MORPHOLOGY AND SYNTAX OF GERUNDS IN TRUKU SEEDIQ
: A THIRD FUNCTION OF AUSTRONESIAN "VOICE" MORPHOLOGY

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR IN PHILOSOPHY

IN

LINGUISTICS

SUMMER 2017

By

Mayumi Oiwa-Bungard

Dissertation Committee:

Robert Blust, chairperson

Yuko Otsuka, chairperson

Lyle Campbell

Shinichiro Fukuda

Li Jiang

Dedicated to the memory of Yudaw Pisaw, a beloved friend

ACKNOWLEDGEMENTS

First and foremost, I would like to express my most profound gratitude to the hospitality and generosity of the many members of the Truku community in the Bsngan and the Qowgan villages that I crossed paths with over the years. I'd like to especially acknowledge my consultants, the late 田信德 (Tian Xin-de), 朱玉茹 (Zhu Yu-ru), 戴秋貴 (Dai Qiu-gui), and 林玉夏 (Lin Yu-xia). Their dedication and passion for the language have been an endless source of inspiration to me. Pastor Dai and Ms. Lin also provided me with what I can call home away from home, and treated me like family.

I am hugely indebted to my committee members. I would like to express special thanks to my two co-chairs and mentors, Dr. Robert Blust and Dr. Yuko Otsuka. Dr. Blust encouraged me to apply for the PhD program, when I was ready to leave academia after receiving my Master's degree. If it wasn't for the gentle push from such a prominent figure in the field, I would never have seen the potential in myself. I could not have completed this dissertation without Dr. Otsuka's countless acute observations and constructive critiques on my analysis and writing. I also owe a great amount of gratitude to her for emotional support and frequent reassurances.

The research for this dissertation was partially funded by a Taiwan Fellowship offered by the Ministry of Foreign Affairs, Republic of China. I am thankful to Dr. Elizabeth Zeitoun and the Institute of Linguistics, Academia Sinica for hosting me for the duration of the fellowship. Dr. Zeitoun's wealth of knowledge in Formosan linguistics and analytical feedback were critical in refining my research questions. She also continues to be supportive and encouraging after the conclusion of the fellowship.

Many current and former students in the University of Hawai'i at Mānoa Department of Linguistics have had significant impact in my life, both at the professional and personal levels. Apay Tang, Jonathan Kuo, Victoria Chen, and Ya-chi Yeh were directly involved with my research in various aspects, from recruiting consultants and arranging my field trips to translating for me. Jonathan and Victoria also frequently gave me valuable insights to my research. I am extremely lucky to have such hard-working and brilliant colleagues as you two.

I also deeply appreciate the friendship of many other students that I shared so much laughter and so many tears with. I would like to extend special thanks to Katie Gao, Kate Hardeman-

Guthrie, Sunju Kim, Claire Stabile, Nozomi Tanaka, and Hiroko Sato – thank you for putting up with me over so many years!

My parents have been a source of constant support and comfort throughout my life. I hope I can finally make them feel (somewhat) at ease with the completion of my Ph.D.

Ryan, I don't have enough words to express how grateful I am for your love, patience, understanding, and unwavering support. Thank you for always believing in me. And let's not forget – this dissertation wouldn't have been complete without all the hard work you put into proofreading it!

ABSTRACT

This dissertation is the first account of gerund constructions in Truku Seediq, an Austronesian language spoken in Taiwan with a Philippine-type voice system. Gerunds are an in-between syntactic category exhibiting both prototypically nominal characteristics and prototypically sentential characteristics. They are event- or state-denoting and derived by a productive nominalization process applied to verbal stems.

The aims of this dissertation are to provide a detailed morphosyntactic analysis of Truku gerunds and to seek the historical origins of their morphology. Careful investigation unveils an array of both prototypically nominal and prototypically sentential features associated with the construction, confirming previous cross-linguistic characterization of gerunds. They also manifest typologically unique patterns, such as an interaction with external possession and limitations on voice alternation.

Gerunds in Truku are derived via highly polysemous morphemes. Markers for indicative gerunds also function as voice markers and thematic nominalizers. An identical pattern of reduplication not only creates subjunctive gerunds but also marks futurity and derives means/manner nominalizations. Despite surface similarities, I demonstrate these functions to be distinct from one another on semantic, distributional, and syntactic grounds.

I propose probable scenarios for the development of Truku gerund markers based on their reconstructed functions in the proto language and cross-linguistic comparisons at the diachronic level. Upon scrutiny, it becomes apparent that Seediq added new functions to two sets of preexisting morphemes: voice markers/ thematic nominalizers (indicative gerund formation) and a pattern of reduplication (subjunctive gerund formation). Both can be traced back to Proto-Austronesian and are widely reflected in its daughter languages. Nevertheless, these innovations are unique to Seediq. They are most likely independent innovations.

In spite of distinct origins, indicative and subjunctive gerund markers underwent comparable lines of change that conspired to expand the verbal paradigm of the language. This change involved loss of thematic orientation and a semantic shift towards denotation of events rather than entities. Simultaneously, the original functions of the morphemes were left intact, yielding an unusual level of morphological polysemy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
ACT	active
ACTR	actor
AN	Austronesian
AV	Actor Voice
CAUS	causative
CLS	clausal
CN	common noun
COM	comitative
CV	Circumstantial Voice
CS	change of state
DAT	dative
DES	desiderative
DIST	distal
DO	direct object
DYN	dynamic
EP	external possession
EPC	external possessor construction
EPIS	epistemic
EMPH	emphatic
EXCL	exclusive
EXIST	existential
FIN	finite
FUT	future
GEN	genitive
GER	gerund
HORT	hortative
IMM	immediate
IMP	imperative
INCL	inclusive
IND	indicative
INDF	indefinite
INT	intensifier
INTR	intransitive
INST	instrumental
IPFV	imperfective
IRR	irrealis
LIG	ligature
LOC	locative
LV	Locative Voice
MAN	means/manner
NAV	non-Actor Voice
NEG	negation
NFIN	nonfinite
NFUT	nonfuture
NMLZ	nominalization
NOM	nominative
NRF	non-referential
OBL	oblique
OV	object voice
PAn	Proto-Austronesian
PART	particle
PASS	passive
PAT	patient
PFV	perfective
PL	plural
PM	possessum
PMP	Proto-Malayo-Polynesian
POSS	possessive
PPN	proper noun
PR	possessor
PROG	progressive
PROX	proximal
PST	past
PV	Patient Voice
Q	question
QUOT	quoted speech
REAL	realis
RECP	reciprocal
RED	reduplicant
REFL	reflexive
REL	relative
RES	result
SBJV	subjunctive
SG	singular
SIM	similative
STAT	stative
SUBJ	subject
SVC	serial verb construction
TEMP	temporal
TOP	topic
TR	transitive
UV	Undergoer Voice
VBLZ	verbalization

CHAPTER 1

INTRODUCTION

1.1. Overview

This dissertation provides the first account of gerund constructions in Truku Seediq (ISO 639-3 identifier: trv), an Austronesian language spoken in Taiwan with a Philippine-type voice system. Gerunds are *event- or state-denoting forms exhibiting both prototypically nominal characteristics and prototypically sentential characteristics.* They are derived by a productive nominalization process applied to verbal stems; thus, the semantic relationship between the verb and its gerundivized form is predictable and transparent.

Though other syntactic items demonstrate differing degrees of nominal and sentential characteristics, they all fall somewhere on a cline between prototypical nouns and prototypical sentences. Thus, such constructions are often best captured in comparison with one another. For instance, English nominalizations fall into three classifications: derived nominals, gerunds, and mixed nominalizations (Chomsky 1970).

(1.1) John refused the offer (proposition)
(1.2) I was upset at John's **refusal** of the offer (derived nominal)
(1.3) I was upset at John('s) **refusing** the offer (gerund)
(1.4) I was upset at John's **refusing** of the offer (mixed nominalization)

The three types differ not only in their morphology, but also in respects of productivity, semantic regularity, and various syntactic characteristics, the last of which is indicative of the presence of an internal nominal structure or the absence thereof. Gerunds in Truku resemble those in English in some regards, and yet demonstrate intriguing patterns that offer implications for the typology of nominalized constructions.

Especially curious is the morphology of Truku gerunds: it is homophonous with some of the voice markers in the language. Thus, it sometimes yields intransitive verbs that may *appear* to be marked for Patient Voice (PV). The PV-marked verb in (1.5) and the gerund in (1.6) both carry the affixes *<n>* and *-an*. In Patient Voice, the Patient of the verb serves as the sentential pivot and is marked by the nominative case marker *ka* in (1.5).

(1.5) **s<n>qit-an**=mu ka qhuni nii (Patient Voice)
cut<PFV>-PV=1SG.GEN NOM tree PROX
'I cut these trees.'

(1.6) ini=ku kla **d<n>hq-an** Iming (gerund)
NEG=1SG.NOM know.AV.NFIN arrive<PFV>-GER Iming.GEN
'I did not know about Iming's arrival (lit. arriving).'

In this dissertation, I distinguish constructions like (1.6) on semantic, distributional, and syntactic grounds from both voice-marked verb forms and other types of nominalization in the language.

This dissertation intends to make contributions both at the descriptive, language-specific level, and at the typological level. First, this study directly adds to the sum of knowledge about Seediq since gerunds have not been mentioned in existing descriptions of the language. Gerunds are a particularly understudied aspect of syntax across Formosan languages, as most of the attention on nominalization has been restricted to lexical or thematic nominalization. Therefore, this dissertation will contribute to the understanding of gerunds in Formosan languages and set the framework for analyzing similar phenomena. Consequently, the analysis presented in this dissertation contributes to the general cross-linguistic typology of gerunds and the phenomenon of nominalization in general.

At the historical level, the morphological reanalysis that gave rise to the gerund-deriving morphology of Truku offers previously unreported instances of historical change on two sets of Austronesian morphology: voice markers / thematic nominalizers and Ca- reduplication. The morphology is of great antiquity dating back to Proto-Austronesian (PAn). As demonstrated above, Truku indicative gerund markers are homophonous with some of the Non-Actor voice markers used in the language: reflexes of PAn *<in> (Patient Voice/perfective), *-en (Patient Voice), and *-an (Locative Voice). Reflexes of these markers, as well as reflexes of *Si- (Instrumental Voice), are widely distributed as both voice markers and lexical/thematic nominalizers in the daughter languages. In particular, verbal forms and nominalized forms in Philippine-type languages are often thought to be homophonous. This observation has resulted in proposals that the canonical predicate in these languages is nominal (De Wolf 1988, Himmelmann 1991, Kaufman 2009, Naylor 1980). An alternative proposal suggests that the original function of these PAn morphemes was nominal derivation, though they have since been reanalyzed as verbal affixes (Starosta, Pawley, and Reid 1982). By assigning the gerund-forming function to these morphemes, I am essentially claiming that Truku has added a "third" function to them, whose products are also categorially ambiguous. I propose that this process involved reanalysis of the affixes as general nominalizers, as well as the loss of their thematic orientation towards the Patient or the Location.

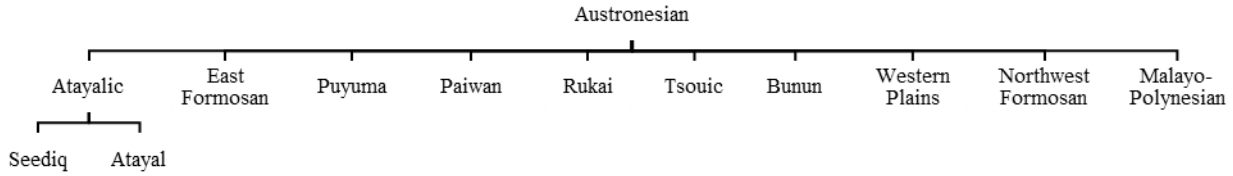
On the other hand, the reduplicative pattern that derives subjunctive gerunds in Truku is shown to have its ultimate roots in PAn *Ca- reduplication, whose functions in the proto-language, according to Blust 1998, included instrument nominalization. I hypothesize that Ca-reduplication gained two other functions in Atayalic: future/irrealis marking and a nominalization strategy that derives the means by which an action is performed, or the manner thereof. I suspect that the gerund-forming function was an extension of either of the two functions, both of which continue to be productive in Truku.

Not only are these proposed historical changes previously unattested, but they curiously seem to conspire to bring about a particular outcome. The indicative and subjunctive sets of gerunds, despite originating in two different historical sources, underwent development with parallel effects. The derived forms came to refer to events or states, rather than particular thematic roles. Simultaneously, the original functions of the morphemes remained productive, which results in homophonous forms with multiple functions. This situation raises questions about the cognitive representation of polysemy and how the mental syntactic representation changes in the process of reanalysis.

1.2. Language profile

Seediq belongs to the Atayalic primary branch of the Austronesian language family (Blust 1999a, figure 1). Truku is commonly thought to be a dialect of Seediq (Tsukida 2009), while some speakers I have interviewed indicated that they consider it an independent language. The term "Formosan languages" is commonly used to refer to the Austronesian languages spoken in Taiwan (with the exception of Yami, which belongs to the Malayo-Polynesian subgroup) without presupposing the existence of a Formosan subgroup. Although there are competing hypotheses on subgrouping within Formosan languages, they agree that Seediq and Atayal form an exclusive subgroup, Atayalic (Blust 1999a, Ferrell 1969, Ross 2009). Other varieties of Seediq include Tgdaya and Toda. The level of mutual intelligibility among the three varieties is unknown at this time. Tgdaya has been described most extensively by Arthur Holmer (1996) and Henry Y.-L. Chang (2000b). In contrast, Toda has not attracted much attention in the literature. By far the most comprehensive description of Truku to date is Naomi Tsukida's (2009) dissertation. Notwithstanding, all dialects are under-documented and their current level of description still leaves much to be desired.

Figure 1.1. Genetic classification of Seediq based on Blust 1999a



Today, Truku is spoken in eastern central Taiwan, predominantly in Hualien County. According to the Council of Indigenous Peoples of Taiwan, ethnic Truku population was approximately 29,555 as of 2014. In 2004, Executive Yuan of the Republic of China recognized Truku as a distinct aboriginal tribe. The distinction reflects the ethnic identity felt among the community, whose members regularly refer to the Tgdaya and the Toda as Seediq, and themselves as Truku. Chiang and Chiang (2005) estimate the speaker population of Truku to be around 7,844, although the precise number is unclear. The difficulty in determining the population size mainly stems from the endangered status of the language; at present, almost no children are naturally acquiring it as their first language. Mandarin Chinese is currently prevalent in Truku households with many adults also possessing knowledge of Taiwanese Hokkien. What remains of the generation educated under Japanese occupation (1895–1945) speaks fluent Japanese. Consequently, their children's generation also has some passive knowledge of Japanese. The level of Truku proficiency among the ethnic group is proportional with age. According to Apay Tang, the results of her psycholinguistic assessment "make it clear that the members of the entire Truku speech community, from ages 10 to 65, are currently on a cline of phonological, lexical, morphological, and syntactic proficiency that decreases with age" (2011:196).

Figure 1.2. Distribution of aboriginal tribes in Taiwan[1]

1.3. Data and methodology

The data used in this dissertation is of two kinds. The data from Truku Seediq and Tgdaya Seediq are both first-hand and second-hand in nature. The data on all other languages were taken from existing literature.

First-hand data was obtained through fieldwork trips I conducted between 2013 and 2016. The fieldwork between July–December 2014 was funded by a Taiwan Fellowship established by the Republic of China Ministry of Foreign Affairs. Data collection that took place during these trips was approved by the University of Hawai‘i at Mānoa Office of Compliance Human Studies Program on June 5, 2012. The study is exempt from federal regulations pertaining to the protection of human research participants under the Code of Federal Regulations 45CRF 46 (2). The case number assigned to the study is CHS #20327 under the title Linguistic Research of the Seediq Language.

Field work for Truku took place in Hualien Country, in the villages of Bsngan (富世) and Qowgan (加灣). My primary language consultants were two native speakers of Truku: a male

[1] Image from Taiwan Indigenous People’s Knowledge Economic Development Association website, http://www.twedance.org/aboriginal00.aspx, Accessed March 6, 2017.

born in 1925, and a female born in 1946. I also consulted other native speakers from time to time, who were born between 1942 and 1951. I conducted field work sessions with the consultant born in 1925 using Japanese, my native language. Mandarin Chinese was used as the intermediary language with all other consultants. I initially used interpreters for Chinese-to-English translation and vice versa, while I started to hold sessions without any interpreters as my Chinese proficiency improved.

A stint of field work for Tgdaya took place in 2016, in the city of Puli, Nantou County. The consultants included a male born in 1954 and a female born in 1952. The contact language was Mandarin Chinese, and Chinese-to-English translation was employed from time-to-time.

All field work sessions for Truku and Tgdaya were audio-recorded for accurate transcription. Whenever in doubt, translations and interpretations were checked with a native speaker of Chinese.

Three methods were employed during data-collection: (i) elicitation, where consultants freely constructed sentences, usually based on a cue given to them, or to provide a translation of a sentence in the contact language, (ii) grammaticality judgment, where I constructed sentences in the target language and asked consultants to judge their acceptability, and (iii) story-telling, where consultants were asked to recite a story in the target language. Each story-telling session was followed up with playbacks of the audio-recording to the consultant in order to facilitate both accurate transcription and sentence-by-sentence translations into the contact language.

Second-hand language data was obtained from existing literature, which includes scholarly publications by various authors, an on-line dictionary, and the Truku Bible.

1.4. Phoneme inventory and orthography

The Truku consonant inventory consists of seventeen phonemes: /p/, /b/, /t/, /d/, /k/, /ɣ/, /q/, /m/, /n/, /ŋ/, /s/, /x/, /χ/, /l̪/, /ʁ/, /w/, and /y/. /t/ and /d/ are palatalized to [ʧ] and [ʤ], respectively, before the high front vowel /i/. Some consonants have word-final allophones; /p/ and /b/ are realized as [k], /m/ as [ŋ], /d/ as [t], /t/ as [t͡s̪], /l̪/ as [ɹ], and /ɣ/ as [w] after /a/ or /u/ or as [y] after /i/.

Table 1.1. Phoneme inventory of Truku consonants

	bilabial	dental/alveolar	palatal	velar	uvular
stop	p b	t d		k	q
fricative		s		x ɣ	χ
nasal	m	n		ŋ	
lateral fricative		ɬ ɮ			
lateral tap		ɺ			
approximant			y	w	

Truku possesses at least four phonemic vowels: high front /i/, high back /u/, low back /a/, and the schwa /ə/ (spelled as *e*), as well as diphthongs /aʉ/, /aɨ/, and /uɨ/. The phonemic status of the mid vowels *e* and *o* has not been established. According to Chiang and Chiang 2005, native speakers variously regard their language as having three, four, (a, i, o, u) or five monophthongal vowels aside from the schwa. They analyze the distribution of [i] and [u] vs. [e/eɨ] and [o/oʉ] to be neither complementary nor free variation. My observations match Chiang and Chiang's. The distributions of [e] and [o] are not entirely predictable, although they frequently appear as allophones of high vowels /i/ and /u/, respectively. Namely, high vowels are lowered when adjacent to a uvular consonant or /ɺ/[2].

Table 1.2. Phoneme inventory of Truku vowels

	front	central	back
high	i		u
mid	(e)	ə	(o)
low			a

The orthography for Truku used in this dissertation is a hybrid of phonetic and phonemic systems. It largely follows the conventions currently in use in the community. Table 1.3 shows the correspondence between the speech sounds and written symbols. Consonants are largely represented phonetically (i.e., allophones are spelled differently). Exceptions to this rule are /ɺ/ and /ɣ/, which are invariably spelled as *r* and *g*, respectively. Contextually conditioned lowering of high vowels is not reflected in the orthography, i.e., they are spelled as *i* and *u*. Where the occurrence of [e/eɨ] and [o/oʉ] is unconditioned, they are written as *ey* and *ow*, respectively (e.g., *m-**ey**dang* 'AV-go missing' and *b**ow**yak* 'wild pig'). The schwa is not written where its existence is predictable; that is, between consonants (e.g., *m-duk* [məduk] 'AV-close': *eduk* [əduk] 'close. AV.NFIN'). An exception to this is homorganic nasal-oral pairs, which are not broken up by

[2] Note that Truku /ɺ/ is a reflex of Proto-Austronesian *R, which Blust (1990) claims to be either an alveolar or uvular trill. This helps explain the effects on adjacent vowels it has in common with uvular consonants in the modern language.

schwa-insertion (e.g., *e<n>duk* [ə**nd**uk] ‘close<PV.PFV>’). Any full vowel that is reduced to the schwa via a morphonological process is treated as a schwa for the purpose of orthographic representation (e.g., *q<m>ita* [qəm**i**ta] ‘see<AV>’: *qta-an* [qə**t**aan] ‘see-PV’).

Table 1.3. Truku orthographic conventions adopted in this dissertation

phone	written as	phone	written as
p	*p*	a	*a*
b	*b*	i	*i*
m	*m*	u	*u*
t, t͡s	*t*	ə	*e*
ʧ	*c*	aɪ̯	*ay*
d	*d*	uɪ̯	*uy*
dʒ	*j*	aʊ̯	*aw*
s	*s*	e/eɪ̯	*ey*
n	*n*	o/oʊ̯	*ow*
l̪ , ɹ	*r*		
ɮ	*l*		
y	*y*		
k	*k*		
x	*x*		
ɣ	*g*		
ŋ	*ng*		
w	*w*		
q	*q*		
χ	*h*		

1.5 Organization of the dissertation

The remainder of this dissertation is organized as follows. Chapter 2 outlines the morphosyntax of Truku Seediq, in order to facilitate a deeper understanding of gerunds, the central theme of the dissertation. The sketch includes such aspects as the language’s voice system, verbal morphology, pronominal system, word order, and various syntactic processes relevant to the discussion of gerunds presented in the chapter to follow. Chapter 3 first illustrates the phenomenon of Truku gerunds in terms of morphology, semantics, distribution, and syntactic behavior. The chapter also contains a sketch of various nominalized constructions in the language as a point of comparison with gerunds. In addition, the chapter is supplemented with an analysis of external possessor constructions, which lends a supporting argument for the gerunds’ nominal properties. Chapter 4 probes gerund-like constructions attested in other Formosan speech varieties. Chapter 5 summarizes the functions of PAn *Ca- reduplication and its reflexes in the daughter languages, with an emphasis on Atayalic varieties. Through cross-linguistic

comparisons, these two chapters aim to find the historical roots of Truku gerundive morphology. Chapter 6 concludes the dissertation by summarizing the findings and analyses of this study, and addresses its limitations and prospects of future research.

CHAPTER 2

MORPHOSYNTACTIC SKETCH OF TRUKU SEEDIQ

This chapter presents basic morphosyntactic characteristics of Truku in order to set the stage for the discussion of gerunds in the language.

2.1. The voice system

Like most other Formosan languages, Truku has a Philippine-type voice system that goes beyond the two-way active vs. passive split. Concretely, verbal morphology corresponds to the thematic role born by the syntactic "pivot" of the sentence, in the sense of Dixon 1979; that is, a language specific, surface "subject." [3] In the Tagalog examples below, the pivot is morphologically marked by *ang*, and the corresponding verbal morphology appears in bold. Non-pivot Actors receive the same morphological marking as genitive.

(2.1) *Tagalog* (Foley and Van Valin 1984:63, glosses mine)

a.	b<**um**>ili	ng	isda	sa	bata	ang	lalake	(Actor Voice)
	buy<AV>	GEN	fish	OBL	child	NOM	man	
	'The man bought some fish from the child.'							
b.	b<**in**>ili	ng	lalake	sa	bata	ang	isda	(Patient Voice)
	buy<PFV.PV>	GEN	man	OBL	child	NOM	fish	
	'The man bought the fish from the child.'							
c.	b<in>ilh-**an**	ng	lalake	ng	isda	ang	bata	(Locative Voice)
	buy<PFV>-LV	GEN	man	GEN	fish	NOM	child	
	'The man bought some fish from the child.'							

Early literature described these systems as having multiple "passives" (Bell 1983, Blake 1925, Bloomfield 1917, Givón 1979). Others treat Philippine-type languages as having a (split-) ergative system (Cooreman, Fox and Givón 1984, Payne 1982). Upon this view, so-called Patient Voice is the default transitive, while Actor Voice is essentially an antipassive construction, and Locative and Circumstantial Voice constitutes applicative constructions. Shibatani (1988) and De Wolf (1988) reject both of these views by concluding that the Philippine system is neither accusative nor ergative. Similarly in Foley's (2007) "symmetrical voice system" approach, no one NP with a particular thematic role is preferred as pivot, and all voice alternations are equally

[3] The notion of the grammatical subject in Philippine-type languages has caused much controversy. Namely, subject properties appear to be divided between the pivot (also called nominative NP/argument, absolutive NP/argument, topic, etc.) and the Actor. While the facts have led some to conclude that Philippine-languages lack the subject (Schachter 1976, Foley and Van Valin 1984), others treat either the nominative argument (Bloomfield 1917, Blake 1925, Bell 1976, Kroeger 1993) or the Actor (Carrier-Duncan 1985, Gerdts 1988, Payne 1982) exclusively as the subject.

marked syntactically. This contrasts with accusative and ergative systems in which one voice type is considered more basic than the other. The situation in Truku is analogous to other Philippine-type languages to the extent that no single voice type is morphologically unmarked. Simultaneously, evidence is inconclusive as to whether PV can be considered canonically transitive in a pragmatic sense. Thus, in this dissertation, I follow the symmetrical voice approach. Following the common convention in Formosan literature, I gloss the syntactic pivot as nominative, and the Actor of a NAV-marked verb (non-pivot Actor) as genitive.

2.2. Verbal morphology

Truku has a four-way Philippine-type voice distinction. In general, each main verb carries verbal morphology that corresponds to the thematic role of the pivot: the nominative argument. There are two sets of voice markers in Truku, finite and nonfinite.

2.2.1. Finite voice morphology

Actor Voice (AV) is indicated by *<m>*, *m-*, or *Ø*. The allomorph selection of Actor Voice is lexically determined (see Tsukida 2009 for a detailed account).

(2.2) s<m>bug=ku kingal bowyak
shoot<AV>=1SG.NOM one wild.pig.OBL
'I shot a wild pig.'

I differentiate between the Actor and the Agent, the former of which does not presuppose agentivity. I treat the Actor as a macrorole in Foley and Van Valin's (1984) sense and proto-role in Dowty's (1991) sense, which extends not only to the Agent, but also to the Experiencer and the Undergoer. Most intransitive verbs are regularly marked with one of the allomorphs of the AV marker, regardless of volitionality.

(2.3) wada m-tucing qsiya ka qtahi
PFV AV-fall water.OBL NOM ant
'The ant fell in the water.'

(2.4) dahuq ka Kuras da
arrive.AV NOM Kuras CS
'Kuras arrived.'

The AV morphology has a variant in the form of *me-*[4] / *mp-* in the irrealis mood. Whereas realis forms denote events that have already taken place, irrealis forms denote events that have yet to

[4] Following the general rule of the language, the realis AV marker *m-* automatically inserts a schwa after it if the verb stem begins with a consonant. On the other hand, the irrealis AV marker has an inherent schwa. Thus, the contrast between the two markers is preserved in cases where the verb starts in a vowel e.g., m-adas /ˈmadas/ 'AV-take/bring' vs. me-adas /məˈadas/ 'AV.IRR-take/bring'.

occur (2.5–6), express prospective aspect (2.7), or express habitual aspect in the present time (2.8). Like its realis counterpart, the allomorphy of this prefix is lexically determined.

(2.5) me-tabug=ku huling saman
AV.IRR-nurture=1SG.NOM dog.OBL tomorrow
'I will feed the dog tomorrow.'

(2.6) mp-griq=ku kamut
AV.IRR-twist=1SG.NOM motorcycle.OBL
'I will drive a motorcycle.'

(2.7) saw mp-k-stuq siida ka tama=mu o,…
SIM IRR-STAT.NFIN-severed when NOM father=1SG.GEN TOP
'When my father was about to die…' (*Soyang Patas,* Genesis 50:5)

(2.8) ini=ku skuxul mp-qhnga ka laqi=su
NEG=1SG.NOM like.AV.NFIN AV.IRR-misbehave NOM child=2SG.GEN
'I do not like that your child is naughty.'

The Patient proto-role also includes the direct Patient, the Goal, and the transported Theme. PV verbs are formed with *-un*, *<n>/n-*, or *-an*. In elicitation, verbs marked with *-un* refer to events that have yet to occur (2.9). However, the suffix cannot be analyzed as a marker of future tense. Not only can it co-occur with a perfective marker (2.10), but it also appears in the perfective context in natural speech.

(2.9) biq-un=ta lqi-an ka patas
give-PV=1PL.INCL.GEN child-OBL NOM book
'We will give the book to a child.'

(2.10) wada=misu paq-un[5]
PFV=1SG.GEN:2SG.NOM hit-PV
'I hit you.'

The primary function of the infix/prefix *<n>/n-* (henceforth simply written as *<n>*) is to indicate perfective aspect. However, it yields a PV reading (2.11) unless it co-occurs with another voice marker (2.12). For this reason, I will treat the affix as a portmanteau PV/perfective marker when occurring on its own.

(2.11) s<n>ipaq=mu ka huling gaga
hit<PFV.PV>=1SG.GEN NOM dog DIST
'I hit that dog.'

[5] In Truku, verb roots are often truncated when suffixation would otherwise yield a word that is trisyllabic or longer e.g., *sipaq* 'hit (root)' vs. *paq-un* 'hit-PV', rather than **sipaq-un*. Most verb roots are disyllabic, and it is the first (C)V that is omitted.

(2.12) s<m><n>ipaq=ku huling
hit<AV><PFV>=1SG.NOM dog.OBL
'I hit a/the dog.'

The suffix *-an* is a reflex of PAn *-an, the Locative Voice (LV) marker, a function that is preserved in Truku (2.13). Aside from its LV function, *-an* also corresponds to the Patient pivot with most verbs (2.14).

(2.13) krut-an qsurux bubu ka puy-an
cut-LV fish.OBL mother.GEN NOM cook-LOC.NMLZ
'Mother cut the fish in the kitchen.'

(2.14) bhang-an=mu ka uyas qbhni
hear-PV=1SG.GEN NOM song bird
'I hear a bird song.'

Circumstantial Voice, marked with *s-*, selects Instrument (2.15), Benefactive (2.16), transferred Theme (2.17), or a reason (2.18) as pivot.

(2.15) s-krut=mu qsurux ka yayu nii
CV-cut=1SG.GEN fish.OBL NOM knife PROX
'I cut fish with this knife.'

(2.16) s-barig=na phpah ka kuyuh=na
CV-buy=3SG.GEN flower.OBL NOM woman=3SG.GEN
'He bought flowers for his wife.'

(2.17) s-bgay Abus Lking-an ka pila
CV-give Abus.GEN Lowking-OBL NOM money
'Abus gives money to Lowking.'

(2.18) s-lingis Kuras ka bubu=na
CV-cry Kuras.GEN NOM mother=3SG.GEN
'Kuras cries for / because of his mother.'

2.2.2. Tense, aspect, and mood

Tense, aspect, and mood are usually encoded via a preverb (2.10, repeated here as 2.19; 2.20). An exception to this tendency is the aforementioned AV irrealis marker *me-/mp-*, perfective affix *<n>*, and the desiderative mood marker *km-* (2.21).

(2.19) wada=misu paq-un
PFV=1SG.GEN:2SG.NOM hit-PV
'I hit you.'

(2.20) gisu=ku m-iing Rubiq
PROG.PROX=1SG.NOM AV-look.for Rubiq.OBL
'I am looking for Rubiq.'

(2.21) km-uq-un=mu bi ka yabas gaga
DES-eat-PV=1SG.GEN INT NOM guava DIST
'I really want to eat those guavas.'

Preverbs are a closed class of items that express aspect, mood, etc. They are found across Formosan languages, and their use can be traced back to PAn. These items have been referred to as auxiliaries or auxiliary verbs in some publications (Starosta, Pawley, and Reid 1982, Zeitoun et al. 1996, among others). However, I draw a distinction between preverbs and auxiliaries in Truku, motivated by the fact that the respective verb form that follows them comes from two different sets. Examples of preverbs in Truku include *hana* 'just recently, finally', *yaa* 'uncertain', *naa* 'should, had better', *ida* 'surely (epistemic)', *ana* 'even, even if', *gisu* 'progressive (proximal)', *wada* 'perfective', and *mha* 'future', to name a few. Preverbs are followed by finite verb forms that are either fully inflected or voice-inflected with no tense or aspect morphology.

(2.22) gisu mp-dahuq ka Kuras da
PROG.PROX AV.IRR-arrive NOM Kuras CS
'Kuras is arriving soon.'

(2.23) hana=mu n-arig ka tlng-an nii
IMM.PST=1SG.GEN PFV.PV-buy NOM sit-LOC.NMLZ PROX
'I just bought this chair.'

On the other hand, verbs that follow auxiliaries are in the nonfinite form, as elaborated on in the next section. Auxuliaries include *ini* 'negative marker', *iya* 'negative imperative', *asi* 'must', and *acih* 'nearly (did)'. An extensive list of preverbs and auxiliaries can be found in Tsukida (2009:421–2).

2.2.3. Nonfinite voice morphology

Aside from the finite voice markers discussed in the previous section, Truku has a second set of voice markers for nonfinite forms. Nonfinite voice morphology has a three-way split, with zero-marking for AV, *-i* for PV and LV, and *-ani* for CV. These forms appear in imperative sentences (2.24–26) and after auxiliary verbs (negation; 2.27–29).

(2.24) imah qsiya!
drink.AV.NFIN water.OBL
'Drink some water!'

(2.25) mah-i (ka)[6] qsiya nii!
drink-PV.NFIN NOM water PROX
'Drink this water!'

(2.26) prq-ani=ku dha balung rudux han
break-CV.NFIN=1SG.NOM two egg.OBL chicken PART
'Break two chicken eggs for me.'

(2.27) ini qbahang kari me-tabug ka seediq gaga
NEG listen.AV.NFIN word ACTR.NMLZ-nurture NOM person DIST
'That person does not listen to the pastor.'

(2.28) ini=mu bhang-i ka uyas qbhni
NEG=1SG.GEN hear-PV.NFIN NOM song bird
'I do not hear the bird song."

(2.29) ini they-i laqi ka sapah nii
NEG play-LV.NFIN child.GEN NOM house PROX
'The child did not play in this house.'

Note that I adopt the term "nonfinite" in a non-traditional sense, to the extent that nonfinite forms serve as primary predicates in imperative sentences. This decision was partially motivated by the lack of a term that better captures the functions of the verb forms. They have previously been dubbed "conegative" by Holmer (1996) and "dependent" by Ross (2002, 2009). The former term can be misleading, given that the forms do not necessarily co-occur with negation. The latter is not any more accurate than "nonfinite" since the forms are not dependent on another element in imperative sentences. An alternative is to analyze the contrast as a realis/irrealis mood distinction. This choice was ruled out for the sake of paradigmatic symmetry. As illustrated in §2.2.5, stative verbs also have two distinct markers: *m-*/*Ø* (finite) and *k-* (nonfinite). While the distribution of *k-* partially overlaps with nonfinite voice markers, *k-* also co-occurs with the perfective marker <*n*> in a number of contexts. Thus, the prefix can hardly be regarded a marker of irrealis mood.

2.2.4. Hortative morphology

Like nonfinite forms, hortative forms have a three-way split of *-a* (AV; 2.30), *-aw/-ay* (PV/LV; 2.31), and *-anay* (CV; 2.32). According to Ross (2009), it is likely the hortative forms were derived from the Proto-Austronesian optative/hortative **-a* followed by the imperative suffixes *-u (PV) and *-i (LV). In any case, it is not clear what conditions regulate the alternation between *-aw* and *-ay* in Truku (Tsukida 2009).

[6] The nominative marker is often optional in NAV imperative sentences.

(2.30) smla-a=ta seediq…
make-HORT.AV=1PL.INCL.NOM person.OBL
'Let us make mankind…' (*Soyang Patas*, Genesis 1.26)

(2.31) p-k-brag-aw/-ay=ta
CAUS-STAT.NFIN-long-HORT.PV/-HORT.PV=1PL.GEN
ka s-sluhay=ta hug
NOM GER.SBJV-study=1PL.GEN Q
'Let us extend our study (time).'

(2.32) py-anay=misu
cook-hort.CV=1SG.GEN:2SG.NOM
saw kuxul uq-un tama=su
SIM like.CV eat-PAT.NMLZ father=2SG.GEN
'Let me prepare some food for you, the way your father likes it.' (*Soyang Patas,* Genesis 27.9)

Truku voice markers are summarized in the table below.

Table 2.1. Truku voice markers

	AV		PV		LV	CV
	realis	irrealis	non-perfective	perfective		
finite	<m>, m-	me-, mp-	-un	<n>, n-	-an	s-
nonfinite	Ø		-i			-ani
hortative	-a		-aw, -ay			-anay

2.2.5. Stative morphology

Stative verbs show different patterns from non-stative verbs in that the former do not normally show any voice morphology. Stative verbs are normally either marked by *m-* (2.33) or zero-marked (2.34) in the finite form.

(2.33) m-dakil ka qhuni nii da
STAT.FIN-mature NOM tree PROX CS
'This tree has grown.'

(2.34) paru ka duriq Iming
STAT.FIN.big NOM eye Iming.GEN
'Iming's eyes are big.'

The nonfinite stative marker *k-* is in complementary distribution with *m-*/*Ø*. Zeitoun and Huang (2000) establish PAn *ma- and *ka- (reflected in Truku as *m-* and *k-*) as markers of stativity in finite and nonfinite contexts, respectively. By nonfinite contexts, the authors mean "causative, imperative, irrealis, and [NAV]," as well as "declarative/imperative, negative, reciprocal, and reflexive constructions" (2000:398). In Truku, *k-* is used in an imperative form (2.35) or when a

stative verb is subsumed under auxiliaries like negative *ini* (2.36) and negative imperative *iya*. These two environments are shared by the nonfinite forms of action verbs (§2.2.3).

(2.35) k-drumut!
STAT.NFIN-diligent
'Be diligent!'

(2.36) ini k-dakil ka qhuni nii nia na
NEG STAT.NFIN-mature NOM tree PROX yet yet
'This tree has not grown yet.'

Unlike nonfinite voice markers, *k-* also appears when a stative verb is causativized (2.37), bears the irrealis marker *mp-* (2.38), or bears NAV morphology as the first verb in a serial verb construction (SVC)[7] (2.39). Causativization is illustrated in §2.7, tense/aspect/mood morphology in §2.2.1, and SVCs in §2.8).

(2.37) p-k-naqih=su kuxul knan
CAUS-STAT.NFIN-bad=2SG.NOM feeling.OBL 1SG.OBL
'You made me sad.'

(2.38) aji mp-k-dakil ka qhuni nii da
NEG IRR-STAT.NFIN-mature NOM tree PROX CS
'This tree will not grow anymore.'

(2.39) wada=na k-hdu-un bi m-ekan
PFV=3SG.GEN STAT.NFIN-finished-PV INT AV-eat
ka kana uq-un
NOM all eat-PAT.NMLZ
'He has finished all of the food.'

2.3. Nominal predicates

Nominals in Truku serve as the predicate with no overt copula. In the predicate position, the first item of a noun phrase hosts pronominal clitics. Therefore, it is the stative verb *malu* 'good' in (2.41) that the first person singular clitic attaches to, rather than the noun *kuyuh* 'woman' that it modifies.

(2.40) kuyuh=ku (ka yaku)
woman=1SG.NOM NOM 1SG
'I am a woman.'

(2.41) malu=ku kuyuh (ka yaku)
good.STAT.FIN=1SG.NOM woman NOM 1SG
'I am a good woman.'

[7] While stative verbs typically do not inflect for voice, SVCs allow them to do so. This is because the verbs constituting an SVC have a single argument structure, with all voice and tense/aspect marking obligatorily on the first verb.

2.4. Existential constructions

There are two distinct existential constructions in Truku: (i) general existential, which indicates existence ('there exists *x*') including possessive existential ('there exists *y*'s *x*' = '*y* has *x*'), and (ii) locative existential 'to be/stay/live (somewhere)'.

General existential is marked by the predicate *niqan* in the affirmative (2.42), and *ungat* in the negative (2.43).

(2.42) niqan kingal ka uwa seuxal msa
EXIST one NOM young.woman before AV.QUOT
'It is said that there was a young woman of old time.'

(2.43) ungat ka hukut s-sbut=mu huling
NEG.EXIST NOM cane CV-beat=1SG.GEN dog.OBL
'The cane with which I beat the dog is gone.'

Note, however, that it is often the possessor, rather than the possessee, of the possessive existential construction that is marked nominative. These *external possessor constructions* are a common phenomenon in Truku and will be discussed in §3.7.

(2.44) niqan laqi=na ka Rubiq
EXIST child=3SG.GEN NOM Rubiq
'Rubiq has a child / children.'

(2.45) ungat laqi=na ka Rubiq
NEG.EXIST child=3SG.GEN NOM Rubiq
'Rubiq has no children.'

Locative existential is expressed via the verb *eniq*, which can be translated as 'to be/stay/live (somewhere)'.

(2.46) a. m-eniq=ku sapah Kuras
AV-LOC.EXIST=1SG.NOM house.OBL Kuras.GEN
b. niq-an=mu ka sapah Kuras
LOC.EXIST-LV=1SG.GEN NOM house Kuras.GEN
'I am at Kuras' house.'

Locative existential is negated with the negative marker *ini*.

(2.47) ini eniq hini ka Kuras
NEG LOC.EXIST.AV.NFIN here NOM Kuras
'Kuras is not here.'

(2.48) ini=na niq-i ka sapah nii
NEG=3SG.GEN LOC.EXIST-LV.NFIN NOM house PROX
'He does not live in this house.'

Although the general existential *niqan* is homophonous with the LV form of *eniq*, and could be analyzed as historically originating from it, I treat it as a lexicalized form. This is because it does not alternate for voice and also lacks the nonfinite form. The pivot of the general existential *niqan/ungat* is either the person or the object whose existence is being predicated, rather than a location (2.42–3). The pivot of *eniq* is the Actor when the verb is inflected for AV (2.46a), and the Location when it is inflected for LV (2.46b). While *eniq* appears in its nonfinite form when negated (2.47–8), negation of a general existential sentence is achieved through the negative existential predicate *ungat* (2.43).

2.5. Word order and case-marking

Truku word order is predicate-initial. There are three types of case: nominative, oblique, and genitive. Nominative NPs are normally in the sentence-final position. Actors in NAV sentences (non-pivot Actor) are marked genitive and follow the verb. All other NPs are marked oblique and also appear post-verbally. The relative ordering of the non-pivot Actor and oblique arguments varies, but it is not free (2.50–51).

(2.49) s<m>ais kingal lukus knan ka bubu=mu
sew<AV> one clothes.OBL 1SG.OBL NOM mother=1SG.GEN
'My mother sewed me a piece of clothing.'

(2.50) a. wada biq-an patas Kuras ka laqi=na
PFV give-PV book.OBL Kuras.GEN NOM child=3SG.GEN
b. *wada biq-an Kuras patas ka laqi=na
PFV give-PV Kuras.GEN book.OBL NOM child=3SG.GEN
'Kuras gave his child a book.'

(2.51) a. b<n>gay Kuras knan ka lukus nii
give<PFV.PV> Kuras.GEN 1SG.OBL NOM clothes PROX
b. *b<n>gay knan Kuras ka lukus nii
give<PFV.PV> 1SG.OBL Kuras.GEN NOM clothes PROX
'Kuras gave me these clothes.'

Nominative NPs are regularly marked by *ka*. Oblique case is optionally marked by *-an* on human NPs (e.g., *laqi* 'child' > *lqi-an* 'child.OBL') and personal names, but is otherwise unmarked. Genitive case on noun phrases is also unmarked. Although these patterns render oblique and genitive cases phonetically identical in many instances, they are formally differentiated in pronominal sets (see §2.6). In contrast, the Tgdaya dialect preserves the genitive case marker *na* before noun phrases (compare 2.52–53).

(2.52) *Truku Seediq*
sapah Pawan
house Pawan.GEN
'Pawan's house'

(2.53) *Tgdaya Seediq* (Holmer 1996:73, gloss mine)
sapah na Pawan
house GEN Pawan
'Pawan's house'

Kuo (2014) hypothesizes that the enclitic *=na* '3SG.GEN' derives from the Proto-Seediq genitive case marker *na, which has been lost in Truku. Incidentally, *na seems to be reflected in the possessive marker *n=* in Truku.

(2.54) n=laqi=mu ka nii!
POSS-child=1SG.GEN NOM PROX
'This is my son's!' (*Soyang Patas*, Genesis 37.33)

There are some exceptions to the correspondence between voice morphology and *ka*-marked NPs. Time expressions like *sayang* 'now' and *saman* 'tomorrow' may appear *ka*-marked in the final position of an AV sentence. In such cases, the Actor is either realized as a nominative enclitic or a topicalized element. Thus, the Actor remains the syntactic pivot, whereas the time expression is only morphologically nominative.

(2.55) m-iing=ku Kuras ka sayang
AV-search=1SG.NOM Kuras.OBL NOM now
'I am looking for Kuras right now.'

(2.56) Hana o, sprang aji me-tabug huling ka saman
Hana TOP intentionally NEG AV.IRR-feed dog.OBL NOM tomorrow
'Hana will intentionally not feed the dog tomorrow (e.g., in order to punish it).'

External possession (cf. §3.7) is another instance in which a mismatch between the *ka*-marked NP and voice morphology occurs. The logical subject of (2.57) is clearly 'Masaw's child'. Nevertheless, it is the possessor ('Masaw') that is marked by *ka*.

(2.57) ga l<m>ingis laqi=na ka Masaw
PROG.DIST cry<AV> child=3SG.GEN NOM Masaw
'Masaw's child is crying.'

Along with these semantic mismatches, some constructions regularly lack nominative or *ka*-marked NPs altogether. These include existential constructions (2.58) and meteorological expressions (2.59).

(2.58) ungat laqi=mu
NEG.EXIST child=1SG.GEN
'I have no kids.'

(2.59) q<m>uyuh
rain<AV>
'It is raining.'

2.6. Pronouns

Truku has three sets of free pronouns: oblique, neutral, and possessive (table 2.2). Note that all oblique pronouns have the *-an* ending homophonous with the oblique case marker *-an*, although the forms are not morphologically analyzable in synchronic terms.

(2.60) s<m><n>ipaq knan ka Rubiq
hit<AV><PFV> 1SG.OBL NOM Rubiq
'Rubiq hit me.'

Neutral pronouns are not inherently case-marked. Thus, they may appear with the nominative case marker *ka* (2.61). They also serve as the predicate in pseudo-cleft sentences (2.62).

(2.61) s<m>trung=ku seediq kiya mn-dha ka yaku
meet<AV>=1SG.NOM person.OBL DIST x.times-two NOM 1SG
'I have met that person twice.'

(2.62) aji yaku ka s<m>bu huling p-huqil.
NEG 1SG NOM shoot<AV> dog.OBL CAUS.AV-die
isu ka p-n-huqil
2SG NOM CAUS.AV-PFV-die
'It is not me that shot and killed the dog. It is you (that killed it).'

All possessive pronouns begin with the phonological sequence *ne-*. Its homophony with the possessive marker *n=* (phonologically both /nə/) briefly discussed in §2.3 suggests a common historical origin. However, the pronouns are no longer morphologically analyzable. For instance, the first person singular form is *(n)naku* (/(nə)naku/, where *ne-* is optional), instead of the expected /nəyaku/ (*n=* 'POSS' + *yaku* '1SG').

(2.63) nnaku ka huling wada qduriq
1SG.POSS NOM dog PFV escape.AV
'The dog that escaped is mine.'

In addition to free pronouns, the language has two sets of pronominal clitics[8]: nominative and genitive. The nominative set is used to mark the syntactic pivot such as the Actor in AV clauses (2.61) and the Patient in PV clauses (2.64).

(2.64) rngag-i=ku en-da-an=su
speak-PV.NFIN=1SG.NOM PFV-do-GER=2SG.GEN
'Tell me what you did.'

Genitive clitics are used to mark the possessor in a possessive phrase (2.65–66).

(2.65) mp-tgsa=mu
ACTR.NMLZ-teach=1SG.GEN
'my teacher'

(2.66) a. [mp-tgsa kari Truku]=mu
ACTR.NMLZ-teach word Truku=1SG.GEN
b. *mp-tgsa=mu kari Truku
ACTR.NMLZ-teach=1SG.GEN word Truku
'My Truku language teacher.'

At the clausal level, the genitive set marks the non-pivot Actor (2.67).

(2.67) bhang-an=mu ka uyas qbhni
hear-PV=1SG.GEN NOM song bird
'I hear the bird song.'

Pronominal clitics that represent arguments are obligatorily in the second position of the clause. In other words, they are cliticized to the first item in the clause, whether it is lexical (e.g., nouns as in (2.68) or verbs as in (2.69)) or functional (auxiliaries as in (2.70) or preverbs as in (2.71)). A few items, including conjunction markers and discoure markers such as *yaasa* 'because' and *kiya do* 'then', do not host pronominal clitics (2.72).

(2.68) seediq=ku malu (ka yaku)
person=1SG.NOM good.STAT.FIN NOM 1SG
'I am a good person.'

(2.69) s<m>bug=ku kingal bowyak
shoot<AV>=1SG.NOM one wild.pig.OBL
'I shot a wild pig.'

[8] These are morphosyntactic clitics in the sense of Anderson 2005: "linguistic [elements] whose position with respect to the other elements of the phrase or clause follows a distinct set of principles, separate from those of the independently motivated syntax of free elements in the language" (2005:31). In the case of Truku, such principles include the obligatory second place position in the clause, obligatory clitic climbing from the embedded clause to the matrix clause, and clitic doubling: co-occurrence of a clitic and its corresponding independent form within a single clause.

(2.70) ini=mu bhang-i ka uyas qbhini
NEG=1SG.GEN hear-PV.NFIN NOM song bird
'I do not hear the bird song.'

(2.71) ma=su tluung bukuy balay
why=2SG.NOM sit.AV behind INT
'Why are you sitting so far back?'

(2.72) yaasa s<m><n>ipaq=su knan
because hit<AV><PFV>=2SG.NOM 1SG.OBL
ni l<m>ingis=ku
and cry<AV>=1SG.NOM
'Because you hit me, I cried.'

Clitic doubling occurs with nominative pronouns, but it is optional (2.61, 2.68).

Table 2.2 summarizes the pronoun inventory of Truku. Note that the nominative and genitive pairs are phonologically distinct only in the first person singular, the third person singular, and the third person plural. All other pairs are homophonous.

Table 2.2. Truku pronoun inventory

	free			clitic	
	neutral	oblique	possessive	nominative	genitive
1SG	yaku	knan	(n)naku	=ku	=mu
1PL.INCL	ita	tanan	(n)nita	=ta	
1PL.EXCL	yami	mnan	(n)nami	=nami	
2SG	isu	sunan	(n)nisu	=su	
2PL	yamu	munan	(n)namu	=namu	
3SG	hiya	hiyaan	nhiya	Ø	=na
3PL	dhiya	dhiyaan	endhiya	Ø	=dha
WH	ima	emaan	(n)nima		

When the pivot and the Actor of an NAV sentence are both pronominal, some combinations result in pronoun clusters, also called "compound pronouns" (Kuo 2014) (table 2.3). There are two things to note about these clusters. On the one hand, some clusters can be simply considered a combination of the nominative and genitive pronouns. For instance, one may analyze *=kuna* '1SG.NOM:3SG.GEN' as deriving from *=ku* '1SG.NOM' and *=na* '3SG.GEN'. Other clusters seem to be a fusion of a genitive clitic and a neutral free pronoun, such that *=misu* '1SG.GEN:2SG.NOM' consists of *=mu* '1SG.GEN' and *isu* '2SG', with phonological reduction of some sort. However, note that for some clusters the nominative component precedes the genitive one, while the order is reversed for others. Although past analyses have approached the issue from both person and case (Chang 1997, Holmer 1996, Ochiai 2009), there is no straightforward explanation for the nonuniform ordering. A historical account for these pronoun clusters is discussed in detail in Kuo

(2014). Second, there are some gaps in the inventory of pronoun clusters. Nominative case is zero-marked on third person, so it cannot form clusters that surface phonetically. Holmer and Billings (2014) point out that any combinations of two first persons or two second persons are ruled out on logical grounds. Also absent are combinations of 1PL.INCL or 1PL.EXCL as either of the two components, and 2SG or 2PL as the other. Holmer and Billings attribute this restriction to the syncretism between the nominative and genitive forms. That is, the forms do not provide any clues as to which of the two components bears which case.

Table 2.3. Truku pronoun clusters

		nominative						
		1SG	1PL.INCL	1PL.EXCL	2SG	2PL	3SG	3PL
genitive	1SG				=misu	=maku		
	1PL.INCL							
	1PL.EXCL							
	2SG	=saku						
	2PL	=kunamu						
	3SG	=kuna	=tana	=namina	=suna	=namuna		
	3PL	=kudha	=tadha	=namidha	=sudha	=namudha		

2.7. Causativization

Truku has morphological causatives, which are formed with the prefix *p-*. In AV clauses, *p-* serves as a portmanteau marker for AV and causativity. That is, the AV marker <*m*> does not co-occur with *p-* (2.73). AV causative sentences select the Causer as pivot, while the Causee is marked oblique.

(2.73) p-qduriq huling ka Kuras
CAUS.AV-escape dog.OBL NOM Kuras
'Kuras let the dog escape.'

When the verb takes any of the NAV markers, *p-* simply marks causativization (2.74–75). The pivot is the Causee, and the Causer is marked genitive.

(2.74) p-lkus-un=na ka laqi
CAUS-clothe-PV=3SG.GEN NOM child
'She will make the child get dressed.'

(2.75) ga p-sa-an kingal patas
PFV.DIST CAUS-go-LV one book.OBL
ka babaw tluwan Kuras
NOM top.of desk Kuras.GEN
'There is a book on Kuras' desk.' / 'Somebody put a book on Kuras' desk.'

In many cases, native speakers prefer to have an overt Causer in NAV sentences. However, the Causer seems to be optional for those expressions that typically appear with causative verb forms, such as *p-sa-an* ‘be put’ (2.75; literally ‘caused to go’, used to describe a situation in which an object exists in a certain location as a result of somebody having put it there), *p-hqil-un* ‘killed’ (CAUS-die-PV), and *p-n-k-dngu* ‘wilted’ (CAUS-PFV.PV-STAT.NFIN-dry).

When a verb is marked by both the perfective <*n*> and the causative *p-*, the optional portmanteau quality of both markers makes the resulting form ambiguous between AV and PV. It appears that, in most cases, the intended meaning is clear from the context and/or the pronominal case-marking. This is because *p-* can either simultaneously mark causativity and AV, or causativity alone (if in combination with another voice marker). By the same token, <*n*> can mark both perfectivity and PV at the same time, or perfectivity alone (if in combination with another voice marker). Thus, in (2.76) and (2.77), the phonologically identical verb form *p-n-huqil* has the Causer (i.e., Walis) and the Causee (i.e., the dog) as pivot, respectively.

(2.76) p-n-huqil babuy ka Walis da
CAUS.AV-PFV-die pig.OBL NOM Walis CS
‘Walis killed a pig.’

(2.77) p-n-huqil=mu s<m>bu ka huling
CAUS-PFV.PV-die=1SG.GEN shoot<AV> NOM dog
‘I shot the dog dead.’

The causative marker attaches to the nonfinite form of stative verbs, which is marked by the nonfinite prefix *k-*.

(2.78) p-k-naqih=su kuxul knan
CAUS.AV-STAT.NFIN-bad=2SG.NOM feelings.OBL 1SG.OBL
‘You made me sad.’

(2.79) p-k-naqih-un=mu kuxul ka laqi=mu
CAUS-STAT.NFIN-bad-PV=1SG.GEN feeling.OBL NOM child=1SG.GEN
‘I made my child sad.’

Interestingly, when a causativized transitive verb is in its PV form, the pivot can be either the Patient of the root verb, or the Causee of the causativized verb. In other words, though voice morphology can take scope over the causative prefix, the reverse is also possible.

(2.80) p-bhrag-an=mu ngiyaw ka huling
CAUS-chase-PV=1SG.GEN cat.OBL NOM dog
‘I made the dog chase the cat.’ (PV scopes over CAUS)
or ‘I made the dog to be chased by the cat.’ (CAUS scopes over PV)

(2.81) p-bhrag-an=mu huling ka ngiyaw
CAUS-chase-PV=1SG.GEN dog.OBL NOM cat
'I made the cat chase the dog.' (PV scopes over CAUS)
or 'I made the cat to be chased by the dog.' (CAUS scopes over PV)

However, the interpretation is clear from the context in most cases. Both readings (*the dog chased the cat* and *the cat chased the dog*) are indeed available for (2.80) and (2.81). Regardless, native speakers have told me that it is most likely the dog that chased the cat in either case, simply because dogs are more easily manipulated by humans than cats.

2.8. Serial verb constructions

The definition of serial verb constructions (SVCs) that I adopt here is *two or more verbs juxtaposed without any intervening coordination marker, complementizer, or NPs.*[9] Together, the first verb (V1) and the second (V2) have a single argument structure.

Whereas V1 can take all types of verbal inflection, V2 is always marked as AV and unmarked for time, aspect, and mood (2.82a). Consequently, the pivot of the entire construction is semantically related to V2 (and only optionally to V1), but its thematic role matches the voice marking of V1.

(2.82) a. diy-un=nami t<m>atak rnaaw ka sawki nii
hold-PV=1PL.EXCL.GEN cut<AV> bush.OBL NOM hatchet PROX
'We use this hatchet to cut the bush.' (Tsukida 2009:697, gloss mine)

b. *diy-un=nami tatak-un rnaaw ka sawki nii
hold-PV=1PL.EXCL.GEN cut-PV bush.OBL NOM hatchet PROX
Intended: 'We use a hatchet to cut the bush.' (Tsukida 2009:715, gloss mine)

(2.83) *m-n-sa brig-un daya ka bawa
AV-PFV-go buy-PV upper.place.OBL NOM bread
Intended: 'He went up to buy some bread.' (Tsukida 2009:715, gloss mine)

Therefore, PV morphology can be borne even by intransitive verbs serving as V1, provided that V2 is transitive.

(2.84) k-n-drmut-an=mu bi p-patas
STAT.NFIN-PFV-diligent-PV=1SG.GEN INT CAUS.AV-study
ka laqi=mu
NOM child=1SG.GEN
'I devotedly made my child study.'

[9] Even though three-verb serialization is also possible, only two-verb serialization is discussed here.

2.9. Negation

All negative markers in Truku occur preverbally, although they take different forms of verbal complements. The two main negators in the language are *ini* and *aji*, which differ in scope. *Ini* is an auxiliary that negates the verb phrase only, such that in (2.85), 'all of my fingers' are 'not swollen'. Being an auxiliary, *ini* must be followed by a nonfinite verb.

(2.85) nii ini k-labu kana
PROG.PROX NEG STAT.NFIN-swollen all
ka tduling baga=mu
NOM digit hand=1SG.GEN
'None of my fingers are swollen.' (lit. 'My fingers are all not swollen.')

Contrastively, *aji* is a preverb that has a scope over an entire proposition. Thus, *aji* in (2.86) indicates that the statement 'all of my fingers are swollen' is false.[10] By virtue of being a preverb, the verb form that follows *aji* is in the finite form.

(2.86) aji ka nii m-labu kana
NEG NOM PROG.PROX STAT.FIN-swollen all
ka tduling baga=mu
NOM digit hand=1SG.GEN
'Not all of my fingers are swollen.' (lit. 'It is not that my fingers are all swollen.')

Ini is used primarily to describe events that did not take place or states that have not materialized (2.87–88). *Aji* is usually used to express the non-occurrence of events and states in the future (2.89–90), as well as to negate permanent states. *Aji* may also negate any proposition, regardless of its aspectual reading.

(2.87) ini=mu qdal-i n-hapuy shiga
NEG=1SG.GEN feed-PV.NFIN RES.NMLZ-cook yesterday
ka huling
NOM dog
'I did not feed the dog yesterday.'

(2.88) ini k-dakil ka qhuni nii nia na
NEG STAT.NFIN-be.mature NOM tree PROX yet yet
'This tree has not grown yet.'

(2.89) aji=ku me-tabug huling ka saman
NEG=1SG.NOM AV.IRR-feed dog.OBL NOM tomorrow
'I will not feed the dog tomorrow.'

(2.90) aji mp-k-dakil ka qhuni nii da
NEG IRR-STAT.NFIN-be.mature NOM tree PROX CS
'This tree will not grow anymore.'

[10] Note that in both cases, *kana* 'all' is a floating quantifier, which can only be launched by the pivot.

In addition to *ini* and *aji*, *ungat* marks negative existential (2.91), and *iya* marks negative imperative (2.92). Note that existential sentences often lack nominative arguments. Like *ini*, *iya* takes nonfinite verb forms as its complement.

(2.91) ungat laqi=mu
NEG.EXIST child=1SG.GEN
'I have no kids.'

(2.92) iya uq-i o!
NEG.IMP eat-PV.NFIN EMPH
'Don't eat (it)!'

2.10. Relative clauses

According to Aldridge 2004, head-initial relative clauses are the type most commonly found across Austronesian languages, while Philippine and Formosan languages have been known to have both the head-final and the head-initial types. Furthermore, some languages including Seediq and Tagalog have the head-internal type (2.93). Relative clauses are bracketed in the examples that follow.

(2.93) *Tgdaya Seediq* (Aldridge 2004:100, glosses mine)
a. sapah [s<n>malu na tama] (head-initial)
house build<PFV> GEN father
b. [s<n>malu na tama] sapah (head-final)
build<PFV> GEN father house
c. [s<n>malu sapah na tama] (head-internal)
build<PFV> house GEN father
'The house that father built'

In Truku, the head-initial type seems to be preferred over the head-final one. Thus, while some constructions allow for either ordering (2.94), others are only marginally acceptable when head-final (2.95).

(2.94) a. wada m-hgliq kana
PFV STAT.FIN-torn all
ka tru lukus [s<n>ais=mu] da
NOM three clothes sew<PFV.PV>=1SG.GEN CS
b. wada m-hgliq kana
PFV STAT.FIN-torn all
ka [s<n>ais=mu] tru lukus da
NOM sew<PFV.PV>=1SG.GEN three clothes CS
'The three pieces of clothing that I sewed are all already torn.'

(2.95) a. n-paq-an ima ka huling [m-huqil] nii?
PFV-hit-PV who.GEN NOM dog AV-die PROX
b. ?n-paq-an ima ka [m-huqil] huling nii?
PFV-hit-PV who.GEN NOM AV-die dog PROX
'Who hit (and killed) this dead dog?'

Head-internal relative clauses are normally left-extraposed with a topic marker. In (2.96), the nominative case marker preceding the head *lukus* 'clothes' indicates that it is internal to the relative clause.

(2.96) [s<n>ais bubu=mu knan ka lukus] o,
sew<PFV.PV> mother.GEN=1SG.GEN 1SG.OBL NOM clothes TOP
wada m-hgliq da
PFV STAT.FIN-torn CS
'The clothes my mother sewed me are already torn.'

Tsukida 2009 provides just one example in which no extraposition takes place. I have not recorded any such examples in my own data.

(2.97) wada qduriq ka [p-n-huqil Rbiq-an ka seediq]
PFV escape.AV NOM CAUS.AV-PFV-die Rubiq-OBL NOM person
'The person who killed Rubiq has escaped.' (Tsukida 2009:690, gloss mine)

Philippine-type languages typically have a restriction by which only the pivot can be relativized. Thus, if the Actor is relativized, the relative clause must have an AV-marked main verb (2.98). Similarly, for the Patient, the Location, or the Instrument to be relativized, the verb must be in the PV form (2.99), the LV form (2.100), and the CV form (2.101), respectively.

(2.98) ga m-s-dara ka huling [q<m>iyut laqi]
PROG.DIST STAT.FIN-VBLZ-blood NOM dog bite<AV> child.OBL
'The dog that bit the boy is bleeding.'

(2.99) ga m-s-dara ka laqi [q<n>iyut-an huling]
PROG.DIST STAT.FIN-VBLZ-blood NOM child bite<PFV>-PV dog.GEN
'The boy that the dog bit is bleeding.'

(2.100) hini ka sapah [n-niq-an=mu]
here NOM house PFV-exist-LV=1SG.GEN
'This is the house where I used to live.'

(2.101) ungat ka huqut [s-sbut=mu huling]
NEG.EXIST NOM cane CV-beat=1SG.GEN dog.OBL
'The cane with which I beat the dog is gone.'

Of particular relevance to the theme of this dissertation is the characteristics of headless relative clauses. They frequently appear as the subject of pseudo-clefts (2.102). Note that the

relative clause is phonologically identical with the predicate in its canonical declarative counterpart, which is not clefted (2.103).

(2.102) tabaku ka [lax-un=mu]. aji sinaw.
tobacco NOM quit-PV=1SG.GEN NEG wine
'What I quit is (smoking) tobacco, not (drinking) alcohol.'

(2.103) lax-un=mu ka tabaku
quit-PV=1SG.GEN NOM tobacco
'I quit (smoking) tobacco.'

Some *wh*-words, namely *ima* 'who' in Truku, are obligatorily fronted. Aldridge (2002) analyzes the corresponding patterns in Tgdaya Seediq as pseudo-clefts with a headless relative clause as the pivot.

(2.104) p-qduriq huling ka Kuras
CAUS.AV-escape dog.OBL NOM Kuras
'Kuras let the dog escape.'

(2.105) ima ka [p-qduriq huling]
who NOM CAUS.AV-escape dog.OBL
'Who let the dog escape?' (lit. 'The one that let the dog escape is who?')

(2.106) *p-qduriq huling ka ima
CAUS.AV-escape dog.OBL NOM who
IM: 'Who let the dog escape?'

When occurring in a possessive phrase, *ima* 'who' pied-pipes the entire phrase to the cleft position.

(2.107) a. huling ima ka wada qduriq
dog who.GEN NOM PFV escape<AV>
b. *wada qduriq ka huling ima
PFV escape<AV> NOM dog who.GEN
c. *ima ka wada qduriq (ka) huling
who NOM PFV escape<AV> NOM dog
'Whose dog escaped?'

On the other hand, *manu* 'what' is only optionally fronted (2.108a). It also has an option to remain in its original position (2.108b).

(2.108) a. manu ka p-n-qduriq Kuras
what NOM CAUS-PV.PFV-escape Kuras
b. p-n-qduriq Kuras ka manu
CAUS-PV.PFV-escape Kuras.GEN NOM what
'What did Kuras let escape?'

These facts of clefting indicate that voice-marked verb forms in Truku can behave as nominals without further morphological modification. The same observation has been made for other Philippine-type languages, leading some to conclude that the canonical declarative sentences in these languages are essentially copular (De Wolf 1988, Himmelmann 1991, Kaufman 2009).

2.11. Topicalization

Topicalization is pervasive in Truku. In topicalization, an argument is preposed to the initial position and followed by the topic marker *o*. The topicalized argument is co-referenced as a pronominal clitic on the predicate. The Actor can be topicalized regardless of its case, while other arguments can only assume the topic position if they are nominative; that is, when their thematic role corresponds to the verbal morphology. Therefore, in the following examples, topicalization of the Actor yields grammatical sentences in both AV and PV (2.109). In contrast, an attempt to topicalize the Patient in an AV sentence renders it unacceptable (2.110). Sentences in (2.111) are non-topicalized versions of (2.109–110).

(2.109) a. Rubiq o, wada s<m>ipaq huling gaga (AV, Actor topicalized)
Rubiq TOP PFV hit<AV> dog.OBL DIST

b. Rubiq o, wada=na paq-un (PV, Actor topicalized)
Rubiq TOP PFV=3SG.GEN hit-PV
ka huling gaga
NOM dog DIST
'Rubiq, she hit that dog.'

(2.110) a. *huling gaga o, (AV, Patient topicalized)
dog DIST TOP
wada s<m>ipaq ka Rubiq
PFV hit<AV> NOM Rubiq

b. huling gaga o, wada paq-un Rubiq (PV, Patient topicalized)
dog dist TOP PFV hit-PV Rubiq.GEN
'That dog, Rubiq hit it.'

(2.111) a. wada s<m>ipaq huling gaga (AV, no topicalization)
PFV hit<AV> dog.OBL DIST
ka Rubiq
NOM Rubiq

b. wada paq-un Rubiq ka huling gaga (PV, no topicalization)
PFV hit-PV Rubiq.GEN NOM dog DIST
'Rubiq hit that dog.'

CHAPTER 3

A DESCRIPTION OF TRUKU GERUNDS

3.1. General characteristics

In this chapter, I will provide a detailed description of gerunds in Truku. I define gerunds as units that are derived from verbs in the syntax and show both nominal and verbal properties. Gerund formation is a productive process in Truku and can be applied to practically any verb. There are two distinct sets of gerund-deriving morphology. I analyze them as signaling indicative and subjunctive moods, based on the semantics and distribution of the resulting forms. As shown in §3.3, while indicative forms imply that an event has already occurred, or is highly likely to occur, the contexts in which subjunctive forms appear are more limited. Aside from appearing in purposive sentences, their presence is predicted by the type of matrix verbs involved: verbs of indirect causation (e.g., *ddug* 'urge', *dakar* 'forbid', *durum* 'request') and expectation (*tga* 'to wait for').

The most unique feature of Truku gerunds is their morphological forms. Namely, the indicative paradigm is morphologically identical to PV marked verbs in the language, taking <*n*>, *-an*, or *-un*, depending on the intended tense/aspect interpretation. Nonetheless, facts about selection of the gerundive subject compel the conclusion that these forms are *voice-neutral*. Specifically, the subject is always the Actor (whether S or A), rather than the Patient. Unlike voice-markers, these morphemes are not oriented towards any particular thematic roles. Thus, the function of <*n*>, *-an*, and *-un* in these instances must be simply to mark gerunds. In fact, voice alternation in gerunds seems to be blocked by the homophony between voice markers and gerund markers. In contrast to indicative gerunds, subjunctive gerunds are formed via Ce- reduplication: reduplication of the first consonant of the verb stem followed by the vowel *e* ([ə]). Though there is nothing in principle that would prevent the subjunctive forms to be voice-marked, they, too, are voice-neutral and have the same subject selection as the indicative forms. The historical relationship between the gerund markers and the voice markers will be explored in Chapter 4.

Gerunds are clausal by nature, and must have an argument structure. Specifically, a gerund requires at least the "subject": the Actor that is marked in genitive case. However, the subject can be phonologically silent for subjunctive gerunds as long as coreference can be retrieved from the context. Semantically, gerunds refer to events or states rather than entities, abstract notions, or propositions. The remainder of this chapter is organized as follows.

In §3.2, in order to better define the constructions in question, I will provide a brief literature review on theoretical approaches to different types of nominalization cross-linguistically. The main focus will be placed on English gerunds and derived nominals. This will also serve to establish a set of criteria to assess verbal vs. nominal features of this in-between category.

Following this, §3.3 will present the two paradigms of gerunds, indicative and subjunctive, and will investigate what syntactic slots each is distributed in.

Based on the criteria determined in §3.2 and interactions with other language-specific syntactic phenomena, §3.4 will provide a detailed description of the behavior of Truku gerunds. I will demonstrate that the verbal/sentential properties of Truku gerunds include (i) aspectual marking, (ii) subject requirement, (iii) licensing of oblique Patients, (iv) modification via serial verb constructions, rather than juxtaposition of a stative verb, and (v) limited ability to combine with demonstratives. On the other hand, nominal properties of gerunds include (i) distributional patterns (nominative argument, oblique argument, predicate, adjunct, and object of a preposition), (ii) the lack of nominative case, (iii) genitive case-marking on the subject, (iv) negation using *aji* rather than *ini*, only the former of which is used to negate nouns, (v) the compatibility with the preposition *asaw* 'because of/for', and (vi) the inability to form temporal clauses with *siida* 'when'. In addition, I will explore optional nominative marking on the gerundive subject as an exception to the aforementioned lack of nominative case. Following an interim summary (§3.5), I will compare Truku gerunds with other types of nominalization in the language in order to clearly set them apart as productive event- or state-denoting nominalizations that take place at the syntactic level (§3.6). As a reference for interested audiences, §3.7 illustrates the phenomenon of external possession that gives rise to exceptional nominative marking in gerunds.

3.2. Literature on nominalization

Gerunds form a category that demonstrates both prototypically nominal and prototypically sentential properties. Taking English as an example, much attention has been given to what Chomsky (1970) calls derived nominals, gerunds, and mixed nominalization.

(3.1) John refused the offer	(proposition)
(3.2) I was upset at John's **refusal** of the offer	(derived nominal)
(3.3) I was upset at John('s) **refusing** the offer	(gerund)
(3.4) I was upset at John's **refusing** of the offer	(mixed nominalization)

Derived nominals consist of an action-denoting root carrying one of a fixed set of suffixes such as *-al, -ion, -ment, -ity, -ism,* and *-ness,* as well as conversion. Gerunds and mixed nominalization are morphologically identical, formed by the action-denoting root and the *-ing* suffix. However, they differ in their syntactic behavior in that the latter shows more nominal properties than the former, as will be discussed below. Abney 1987 further divides gerunds into Acc(usative)-ing and Poss(essive)-ing types, depending on whether the subject surfaces as accusative or Saxon genitive. This is instantiated by the absence / presence of *'s* on *John* in (3.3).

These distinctions have been captured via varying terminology in the linguistic literature. Abney (1987) refers to Chomsky's "mixed nominalization" as "Ing-of gerunds." They are also often called "nominal gerunds" (e.g., Malouf 1998, Wasow and Roeper 1972), as opposed to "verbal gerunds," the latter of which correspond to Acc-ing and Poss-ing gerunds. In the following sections, I will adopt Chomsky's terminology and reserve the term *gerunds* for instances such as (3.3). This excludes Ing-of types from the category of gerunds, which I will call *mixed nominalization*.

In order to elucidate an array of nominal and verbal characteristics to be sought in Truku gerunds, I will delve into the contrast between English derived nominals and gerunds, as well as mixed nominalizations, in the sections to follow.

3.2.1. Derived nominals

Nominals (including derived, affixed ones as well as underived, unaffixed ones) are typically described as having an optional argument structure. In other words, their complements are optional. However, Jane Grimshaw's (1990) seminal work reinterprets this "optionality" as evidence for two distinct categories of nominals. Namely, *result nominals* lack an argument structure, while *complex event nominals* have arguments that are obligatorily realized in the syntax. In (3.5–6), like their respective base verb, neither *expression* nor *assignment* may surface without a nominal complement.

(3.5) a. We express *(our feelings)
b. The frequent expression of one's feelings is desirable.
c. *The frequent expression is desirable.

(3.6) a. We constantly assign *(unsolvable problems)
b. The constant assignment of unsolvable problems is to be avoided.
c. *The constant assignment is to be avoided.

However, many nominals appear ambiguous as to the presence of an argument. Consider:

(3.7) The expression is desirable.
(3.8) The assignment is to be avoided.

Grimshaw treats *expression* in (3.7) and *assignment* in (3.8) as result nominals, which refer to more or less concrete entities rather than events. These are homophonous yet separate lexical items from those in (3.5) and (3.6).

Aside from the complement's optionality, the genitive element for result vs. complex event nominals obtain different interpretations. *John* in (3.9) can be the possessor, the author, or the taker of the examination, whereas *John* in (3.10) is distinctively more subject-like in that he must be interpreted as the agent of the action.

(3.9) John's examination was long. (result nominal)
(3.10) John's examination of the patients took a long time. (complex event nominal)

Grimshaw also introduces a third category of nominals that include items like *race*, *trip*, *process*, and *event*, which she dubs *simple event nominals*. They also refer to events, and yet lack an argument structure, and behave like result nominals in many ways. In sum, it is not that nominals *optionally* have an argument structure. Rather, complex event nominals have an *obligatory* argument structure that is listed in their lexical entries, while other nominals lack it altogether. Oftentimes, homophonous pairs exist in both of the categories, creating an illusion of the arguments being optional.

3.2.2. Gerunds

On Poss-ing gerunds, Abney (1987:13) remarks as follows: "What makes this construction so perplexing is that it seems to be neither fish nor fowl…On the one hand, it is obviously a sentence; but on the other hand, it is obviously a noun phrase." He claims that the external distribution of gerunds is exactly like that of a noun phrase, referring particularly to positions that "sentences" (*that* clauses) are banned from. In English, this includes the subject in subject-auxiliary inversion (3.11), the embedded subject position (3.12), and the object of preposition (3.13) (Abney 1987:13).

(3.11) a. *Did [that John built a spaceship] upset you?
b. Did [John] upset you?
c. Did [John's building a spaceship] upset you?
(3.12) a. *I wondered if [that John built a spaceship] has upset you.
b. I wondered if [John] has upset you.
c. I wondered if [John's building a spaceship] has upset you.

(3.13) a. *I told you about [that John built a spaceship].
b. I told you about [John].
c. I told you about [John's building a spaceship].

Abney points out that there are a number of differences between Acc-ing gerunds and Poss-ing gerunds, demonstrating that the former are more sentential and less nominal than the latter. Though I will not repeat the whole range of these differences here, two such examples are presented below to illustrate the point. First, when conjoined, *that*-clauses and Acc-ing gerunds do not allow plural agreement on the verb (*bother* in 3.14), whereas Poss-ing gerunds do.

(3.14) a. That John came and that Mary left *bother / bothers me. (*that*-clause)
b. John coming and Mary leaving *bother / bothers me. (Acc-ing gerund)
c. John's coming and Mary's leaving bother / bothers me. (Poss-ing gerund)

Second, Abney maintains that inanimate subjects are unnatural with Poss-ing gerunds. Nevertheless, many native speakers seem to accept examples like (3.15b).

(3.15) a. We were very upset at the refrigerator tipping over. (Acc-ing gerund)
b. ?We were very upset at the refrigerator's tipping over. (Poss-ing gerund)

3.2.3. Derived nominals vs. gerunds

In his highly influential *Remarks on nominalization*, Noam Chomsky (1970) raises three key contrasting aspects between derived nominals and gerunds: (i) productivity, (ii) regularity of semantic relationships with associated propositions, and (iii) whether or not their internal structure identifies them as noun phrases. First, not all propositions have corresponding derived nominals. Conversely, gerunds can be formed on essentially any proposition.

(3.16) a. John is easy / difficult to please.
b. *John's easiness / difficulty to please
c. John('s) being easy / being difficult to please

Second, the relationship between derived nominals and their associated propositions are not regular and quite often idiosyncratic. For instance, while a *fishery* refers to a place where fish are bred, *mockery* can refer to either the action of mocking something/someone, or the target of such an action. In contrast, gerunds hold transparent relations with propositions.

Lastly, the internal structure of derived nominals is characteristically that of noun phrases. Gerunds display a stark contrast with them in each of the following features exemplified below[11]. In these examples, (a)-sentences and (b)-sentences represent derived nominals and gerunds,

[11] The examples were taken from Chomsky 1970, and I have added the ungrammatical counterpart to complete each pair where it was not provided in the original work.

respectively. As shown in the alternation between (3.17) and (3.18), the genitive subject of a derived nominal can be replaced by a determiner.

(3.17) a. John's proof of the theorem
b. John('s) proving the theorem

(3.18) a. the proof of the theorem
b. *the proving the theorem

A derived nominal, but not a gerund can be modified using an adjective (3.19).

(3.19) a. John's unmotivated criticism of the book
b. *John('s) unmotivated criticizing the book

Gerunds in the perfect form can be formed with the auxiliary *have*. Derived nominals lack forms parallel to this.

(3.20) John('s) having criticized the book

Many derived nominals can also be pluralized, whereas gerunds cannot.

(3.21) a. John's three proofs of the theorem
b. *John('s) three provings the theorem

Abney (1987) further points out that gerunds diverge from derived nominals in the following ways. Gerunds are able to assign accusative case to the object while derived nominals require *of* to introduce an object-like element (3.22).

(3.22) a. John's destruction *(of) the spaceship
b. John's destroying the spaceship

Gerunds allow raising to subject (3.23), while derived nominals do not. Verbs including *appear* and *seem* are raising-to-subject verbs. When they take a *to*-infinitive clause as a complement, the subject of the complement clause obligatorily "raises" to the subject of the matrix clause position. This raising operation is schematized in (3.24).

(3.23) a. *John's appearance to be dead
b. John's appearing to be dead

(3.24) a. appears [John to be dead] (pre-raising)
b. John appears [___ to be dead] (post-raising)

Similarly, derived nominals do not permit exceptional case-marking (ECM), whereas gerunds do (3.25). *To*-infinitives typically lack an overt subject since they are unable to assign nominative case. However, ECM verbs assign accusative case to the subject of *to*-infinitive complements, thus allowing the subject to surface in that position (3.26).

(3.25) a. *John's belief Bill to be Caesar Augustus
b. John's believing Bill to be Caesar Augustus

(3.26) John believes [Bill to be Caesar Augustus].

Double objects are acceptable with gerunds, but not with derived nominals (3.27), even with *of*-insertion.

(3.27) a. *John's gift/rental (of) Mary (of) a Fiat
b. John's giving/renting Mary a Fiat

Finally, derived nominals and gerunds behave differently with respect to expressions consisting of a verb and a particle, such as *explain away* and *give up*. In a verbal construction, the particle can be postposed to the position after the object (3.28). Neither position is available for derived nominals (3.29a), whereas both are available within gerunds (3.29b).

(3.28) John explained (away) the problem (away).

(3.29) a. *John's explanation (away) of the problem (away)
b. John's explaining (away) the problem (away)

3.2.4. Mixed nominalization

Mixed nominalizations (also called Ing-of gerunds and nominal gerunds) are *ing*-suffixed forms whose complement is introduced by the preposition *of*. As the label suggests, mixed nominalizations have properties overlapping with both derived nominals and gerunds. Even though all three categories are of in-between-category nature, derived nominals are considered the most noun-like, gerunds the least, and mixed nominalization somewhere in the middle of the continuum.

Wasow and Roeper (1972:45) list an array of characteristics that set mixed nominalization apart from gerunds. Unlike mixed nominalization, gerunds cannot co-occur with articles (3.30–31). Whereas mixed nominalization is modified via adjectives, gerunds are modified via adverbs (3.32–33). Mixed nominalization may be pluralized, while gerunds may not (3.34). The auxiliary *have* can be used to create the perfect form of a gerund, but not of a mixed nominalization (3.35). Lastly, mixed nominalization is negated via *no*, while gerunds are negated via *not* (3.36–37).[12] In each pair, (a) is a mixed nominalization, and (b) is a gerund.

(3.30) a. John enjoyed a reading of *The Bald Soprano.*
b. John enjoyed (*a) reading *The Bald Soprano.*

(3.31) a. The killing of his dog upset John.
b. (*The) killing his dog upset John.

[12] These examples have been modified from Wasow and Roeper 1972.

(3.32) a. I detest loud singing of the national anthem.
b. *I detest loud singing the national anthem.

(3.33) a. *I detest singing loudly of the national anthem.
b. I detest singing the national anthem loudly.

(3.34) a. Sightings of UFO's make Mary nervous.
b. *Sightings UFO's make Mary nervous.

(3.35) a. *John's having reviewed of the article.
b. John's having reviewed the article.

(3.36) a. No reading of Shakespeare is acceptable for John.
b. *No reading Shakespeare is acceptable for John.

(3.37) a. *Not reading of Shakespeare is acceptable for John.
b. Not reading Shakespeare is acceptable for John.

Kratzer (1996) also adds that the Saxon genitive in a mixed nominalization need not be the subject. Therefore, (3.38a) may have an interpretation where Mary is the one that reads Pride and Prejudice, but it is also possible that Mary attends an event where somebody else reads the book. Conversely, the only interpretation available for the gerund (3.38b) is one where Mary is the reader.

(3.38) a. Mary's reading of *Pride and Prejudice*
b. Mary's reading Pride and Prejudice

As noted by Chomsky (1970) and further explored in Harley and Noyer 1998, mixed nominalizations do not allow particle shift in verb-particle complex predicates (3.40a) while gerunds do (3.40b).

(3.39) a. Chris's writing **up** of the paper surprised Pat. (no particle shift)
b. Chris writing **up** the paper surprised Pat. (no particle shift)

(3.40) a. ∗Chris's writing of the paper **up** surprised Pat (particle shift)
b. Chris writing the paper **up** surprised Pat. (particle shift)

In addition to the presence of *of* on the complement, all of the features of mixed nominalizations exemplified above—except for particle shift, since derived nominals do not co-occur with particles, as in (3.29a)—align them with derived nominals rather than gerunds. Nevertheless, mixed nominalizations display some sentential characteristics that do not exist in derived nominals. For instance, the verb's argument structure seems to be preserved in mixed nominalizations (3.41).

(3.41) a. The doctor examined the patient.
b. The doctor's examination (derived nominal)
c. The doctor's examining *(of the patients) (mixed nominalization)

Furthermore, despite Wasow and Roeper's (1972) observation that mixed nominals are modified via adjectives, Jespersen (1940:109) notes cases in which they are modified via adverbs (*regularly* in (3.42) and *aloud* in (3.43)).

(3.42) The shutting of the gates regularly at ten o'clock had rendered our residence very irksome to me.

(3.43) From the daily reading of the Bible aloud to his mother…

In his *Remarks*, Chomsky (1970) notes the limited productivity of mixed nominalization. Indeed, they cannot be formed on all verbs. Abney (1987:80) states that psych verbs have various levels of acceptability in mixed nominalization. The stronger the sense of causation a verb has, the more acceptable it is.

(3.44) Mary's frightening of John
(3.45) Mary's angering of John
(3.46) ?*Mary's amusing of John
(3.47) ?*Mary's boring of John
(3.48) *Mary's liking of John
(3.49) *Mary's hating of John

Furthermore, mixed nominalization interacts with lexical aspect (Aktionsart). According to Borer (2005:239), mixed nominalization consistently obtains an atelic interpretation. Thus, it can be formed on atelic verbs (3.50–52), but not on telic verbs (3.53–55).[13]

(3.50) the sinking of the ship
(3.51) the falling of the stock prices
(3.52) the jumping of the cows
(3.53) *the arriving of the train
(3.54) *the erupting of Vesuvius
(3.55) *the exploding of the balloon

Gerunds are fully productive and are unrestrained by lexical aspect:

(3.56) the ship('s) sinking
(3.57) the stock price('s) falling
(3.58) the train('s) arriving
(3.59) the balloon('s) exploding

[13] Note that all of the examples provided by Borer consist of intransitive verbs, so that the complement introduced by *of* is interpreted as the subject rather than the object. On the other hand, telic transitive verbs, whose object is always coded as the *of* complement, lack this anti-telicity effect (e.g., "The killing of the cat," and "Joe's breaking of the vase" are acceptable).

Decades after *Remarks*, mixed nominalizations seem to remain somewhat of a mystery. However, it is evident that they share common features with both derived nominals and gerunds. Thus, nominality vs. sententiality can be considered as a cline rather than all-or-nothing concepts.

With these general properties of English nominalizations as a foundation, I will describe the in-between characteristics of Truku gerunds in §3.3–3.4. In addition to gerunds, §3.6 will survey morphological and syntactic traits of several other types of nominalization in Truku.

3.2.5. "Lexical nominalization" vs. "syntactic nominalization"

In their typological study of nominalization, Comrie and Thompson (1985:392) argue that, in lexical nominalization, the verb itself "become[s] a noun." In contrast, in clausal nominalization, the verb itself "typically has no nominal characteristics and often has such verbal characteristics as person and number, though it may be lacking in tense-aspect marking," though "clauses undergo certain modifications which allow them to function as noun phrases." From a formal syntactic perspective, the level of nominal vs. sentential characteristics may correspond to the level of projection that has been nominalized. For instance, Abney (1987) equates Poss-ing, Acc-ing, and Ing-of gerunds with the nominalizing affix *-ing* attaching to IPs, VPs, and V, respectively.

However, exactly how do we prove the *locus* of a given instance of nominalization? Although the idea that gerund formation takes place in the syntax is an uncontested one, the same can hardly be said of derived nominals; mixed nominals are even less clear in this regard. The key to this discussion is the notion of *argument strucutre*.

One approach is to conceptualize derived nominals as being listed in the lexicon as entries independent of their base verbs. In §3.2.1, I provided a brief overview of Jane Grimshaw's account on result vs. complex event nominals. In her view, both types of derived nominals are listed in the lexicon independently of their root verbs. Complex event nominals' argument specifications are listed in their own entry, and as such independent of their root verbs' argument structures. Works in the generative tradition further postulate word-derivation mechanisms of sorts to account for the formation of derived nominals from their base verbs. I refer readers to Aronoff 1976, Jackendoff 1975, and Lieber 1980, among others.

Alternatively, the argument structure associated with some nominals is taken to emanate from their internal verbal projection (Alexiadou 2001; Borer 2003, 2005; Harley 2009). In Borer's view, for instance, argument structures are not pre-determined by the information stored in

lexical entries. Rather, "the syntactic structure gives rise to [an event structure] template, or a series of templates, which … determine the interpretation of arguments" (Borer 2003:33). In her theory, information concerning linguistic sound-meaning mapping is stored as *Encyclopedic Items* (*EI*s) that are void of categories and argument structures. An array of EIs merge with grammatical formatives which in a sense, nominalize or verbalize them. Complex event nominals (as well as gerunds) are formed when the nominalizing affix (e.g., *-al*, *-ness*, *-ity*) introduces a nominal structure that dominates a verbal structure. This is schematized as follows (Borer 2003:42):

(3.60) [$_{N}$ -tion/-ing [$_{EP/ASPQ}$ [$_{L\text{-}D}$…L .]]][14]

(3.61) is an instantiation of (3.60) (Borer 2003:51).

(3.61) [$_{NP}$ -tion $_{NOM}$ /-ing $_{NOM}$ [$_{EP}$ Kim [$_{ASPQ}$ the vase [$_{L\text{-}D}$ break/destroy]]]]
'Kim's breaking/destruction of the vase'

In turn, result nominals lack an argument structure because they contain no verbal structure. The nominalizing affix directly dominates the EI within the lexical domain. This kind of approach to the argument structure relies heavily on the syntax, and in turn, reduces the explanatory burden on the lexicon.

For linguists working on Formosan languages, teasing apart one type of nominalization from another presents a challenge due to their shared morphology. Examples (3.62–63) come from the same issue of *Language and Linguistics* (Vol 3, Issue 2), which is dedicated exclusively to nominalizations in Formosan languages. Though many of the authors contributing to the issue draw a distinction between "lexical nominalization" and "syntactic/clausal nominalization," no clear definition of such phenomena is offered. It can be inferred that "lexical nominalization" refers to those nominals derived in the lexicon, and "syntactic/clausal nominalization" to those derived in the syntax.

On nominalizations in the Mantauran variety of Rukai, Zeitoun (2002:273) notes that "the same morphological processes are actually used to derive nominals both through lexical and syntactic processes." In the following examples, the same prefix *a-* is used to form a "derived action nominal" (3.62a) and a "clausal nominal" (3.62b). Zeitoun equates the former as lexical nominalization and the latter as clausal nominalization, per the aforementioned distinction made in Comrie and Thompson 1985.

[14] L-D stands for lexical domain, AspQ for quantity aspect, and EP for event phrase.

(3.62) *Mantauran Rukai* (Zeitoun 2002:273, glosses mine)

a. kala-ʔaom-aə **a**-pa-sola-solatə
NMLZ-Japanese-NMLZ NMLZ-CAUS-RED-DYN.NFIN:study
'Japanese education'

b. ðonaʔi **a**-pa-sola-solatə-ð-inamə...
that NMLZ-CAUS-RED-DYN.NFIN:study-3SG.GEN-1PL.EXCL.OBL
'At that time, when they taught us...'

In a comparable manner, Chang and Lee (2002:365) remark on Kavalan: "The distinction between lexical nominalization and syntactic nominalization is not morphologically realized." The suffix *-an* is used to form "patient nouns" (e.g., *qan-an* 'food (eat-NMLZ)') and "location nouns" (e.g., *uzis-an* 'bathroom (bathe-NMLZ)'), as well as "agentive nouns", as part of the "affixal complex" *pa-…an* (e.g., *pa-Ribaut-an* 'fisher (NMLZ-fish-NMLZ)'). In addition, *-an* is also used to form "gerunds" (3.63).

(3.63) *Kavalan* (Chang and Lee 2002:365, gloss mine)

qaqaytisan zaqis-**an**-na tu paRin
dangerous climb-NMLZ-3SG.GEN ACC tree
'His climbing trees is dangerous.'

As I demonstrate in §3.6, Truku is no stranger to this morphological complication.

The theoretical issue of the locus of various types of nominalization is beyond the scope of this dissertation. What is of relevance here is not whether a certain type of nominal is a product of "lexical" vs. "syntactic" nominalization, but the extent of their productivity, presence/absence of the argument structure, as well as various nominal and verbal/sentential features they exhibit in comparison to one another. Although gerund formation in Truku is evidently an instance of "syntactic" nominalization to the extent that it is applied freely and productively, its argument structure does not *a priori* constitute proof that other types of nominalization are "lexical." Accordingly, I will not claim that any of the other types of nominalization in the language are unequivocally derived either in the lexicon or in the syntax.

3.3. Morphology, distribution, and semantics

This section introduces the paradigm of Truku gerunds, which consists of indicative and subjunctive sets. Building upon these fundamental characteristics of gerunds, §3.4 will characterize their various syntactic aspects.

Major differences between the two sets of gerunds reside in the syntactic slots the forms are distributed in, as well as the kind of semantics they typically encode. Both indicative and subjunctive gerunds denote events or states. More specifically, indicative forms represent known

states as well as events that have already occurred, or are highly likely to occur in the future. On the other hand, subjunctive forms denote hypothetical events or states. Both types occur in a variety of syntactic contexts: as a nominative pivot, an oblique argument, a predicate, an adjunct, and as object of a preposition. However, subjunctive gerunds contrast with indicative ones in their frequent co-occurrence with verbs of indirect causation (e.g., *ddug* 'urge', *dakar* 'forbid', *durum* 'request'), speech act verbs (e.g., *rngag* 'speak (to)'), and verbs of expectation (e.g., *tga* 'to wait for'). Truku gerunds are fully productive, permitting all types of verbs, whether transitive, unergative, unaccusative, or stative, in both indicative and subjunctive mood.

3.3.1. Morphology

There are two sets of morphemes that are used to create gerunds. Indicative gerunds are formed with *-an* in the non-future tense. In most cases, indicative forms carry both *-an* and the perfective infix <*n*>. In the future tense, they are formed with *-un*. Subjunctive gerunds have no aspect or tense specifications and are uniformly derived via Ce- reduplication of the stem.

Table 3.1. Truku gerund paradigm

indicative			subjunctive
non-future		future	
perfective	non-perfective		
<n>…an	*-an*	*-un*	Ce- reduplication

3.3.1.1. Indicative mood

Morphological forms of gerunds in indicative mood depend on two factors: (i) whether the verb is stative or non-stative, and in the latter case, (ii) a tense/aspectual distinction. For non-stative verbs, gerunds derived using the combination of *-an* and the perfective marker <*n*> have a perfective reading (3.64). On the other hand, forms suffixed with *-un* typically denote events that have not occurred (3.65). In both (3.64) and (3.65), the gerund is in the pivot position, following the nominative marker *ka*, while the time phrase 'one hour' serves as predicate. In the examples to follow, gerunds are underlined.

(3.64) kingal iyax tuki ka <u>n-ksa-an=mu</u>
one interval hour NOM PFV-walk-GER=1SG.GEN
'I walked for an hour.' (lit. 'My walking is an hour.')

(3.65) kingal iyax tuki ka <u>tlam-un=mu</u>
one interval hour NOM run-GER=1SG.GEN
'I will run for an hour.' ' (lit. 'My running is an hour.')

Albeit infrequent, forms marked only with *-an*, without the perfective marker, are attested as in (3.66). When they do occur, they seem to have a non-perfective reading. However, it is difficult to analyze these forms with any consistency due to the insufficiency of data.

(3.66) kn-kingal jiyax o, kingal iyax tuki ka tlam-an=mu
every-one day TOP one interval hour NOM run-GER=1SG.GEN
'Every day, I run for an hour.' (lit. 'Every day, my running is one hour.')

Unlike non-stative verbs, stative verbs seem to lack the tense/aspect distinction. Indicative gerunds are almost uniformly derived by adding *<n>…-an* to the stem; that is, the nonfinite stative marker *k-* + root. There are a few lexical exceptions in which the affix appears to lack the initial component *<n>* (*k-nrx-an*: *narux* 'be sick', *k-nqih-an*: *k-naqih* 'be bad', *k-hda-an*: *k-hada* 'be ripe, fully cooked'[15]).

(3.67) bsiyaq ka k-n-pig-an phpah nii da
long.STAT.FIN NOM STAT.NFIN-PFV-wrinkled-GER flower.GEN PROX CS
'These flowers have been dead for a long time.' (lit. 'These flowers' being dead is long.')

I have come across one instance of a possible stative gerund affixed with *-un*. However, not only does it fail to obtain a future tense reading, but its semantics is also indistinguishable from its *(<n>…)-an* counterpart.

(3.68) tru idas ka k-nrx-un=mu da
three month NOM STAT.NFIN-sick-GER?=1SG.GEN CS
'I have been sick for three months.' (lit. 'My being sick is three months.')

(3.69) bsiyaq bi ka k-nrx-an=mu da
long.STAT.FIN INT NOM STAT.NFIN-sick-GER=1SG.GEN CS
'I have been sick for a long time.' (lit. 'My being sick is long.')

Note that the indicative paradigm is morphologically identical to some of PV forms in the language, taking *<n>…-an* or *-un*. Nonetheless, I argue that these forms, as well as subjunctive gerunds, are best analyzed as being unmarked for voice (§3.4.2).

3.3.1.2. Subjunctive mood

Subjunctive gerunds are formed by a type of reduplication. The reduplicant consists of the first consonant of the stem followed by the schwa (e.g., *geuguy* 'steal': *g-geuguy* /ɣəɣə'uɣuy/). If the stem begins in a vowel, only a schwa makes up the reduplicant (e.g., *iyah* 'come': *e-iyah* /ə'iyaχ/). The base for reduplication is the verbal stem rather than the root. A stem may include

[15] Some of these exceptions may be a result of phonological reduction due to two /n/'s otherwise occurring next to each other (e.g., *k-n-(n)rx-an*).

the causative prefix *p-* or the nonfinite stative marker *k-*. In such cases, the reduplicant consists of the /pə/ and /kə/ sequences, respectively.[16]

(3.70) d<m>udug knan g-geuguy=mu pila bubu=mu
urge<AV> 1SG.OBL SBJV.GER-steal=1SG.GEN money.OBL mother=1SG.GEN
ka Kuras
NOM Kuras
'Kuras urged me to steal my mother's money.' (lit. 'Kuras pushed me for my stealing my mother's money.')

(3.71) tga-un=mu ka e-iyah=su
wait-PV=1SG.GEN NOM SBJV.GER-come=2SG.GEN
'I am waiting for you to come.' (lit. 'I am waiting for your coming.')

(3.72) d<m>udug knan ka Iming
urge<AV> 1SG.OBL NOM Iming
p-p-hadut taxa hii knan
SBJV.GER-CAUS-send other body.GEN 1SG.OBL
'Iming urged me to have somebody else take me (somewhere).' (lit. 'Iming pushed me for somebody else's taking me.')

(3.73) k-k-brax hii=mu ka t<m>alang=ku
SBJV.GER-STAT.NFIN-strong body.GEN=1SG.GEN NOM run<AV>=1SG.NOM
'I run for my health.' (lit. '(The reason why) I run is my body's being strong.')

I propose that this pattern of reduplication is a retention of PAn *Ca- reduplication as proposed by Blust (1998) (see Chapter 4). However, since all pre-penultimate vowels lenite to the schwa in Truku, the underlying identity of the reduplicated vowel is opaque.

3.3.2. Semantics

3.3.2.1. Indicative mood

Gerunds denote either events or states. The indicative gerunds in (3.74) and (3.75) denote the state of the flower being dead and the event of my breaking his toy, respectively. In (3.76), the gerund can either mean the fact that Iming arrived or the circumstances under which Iming arrived. By virtue of being a proposition, the sentential complement *dahuq ka Iming* 'Iming arrived' in (3.77) may only refer to the fact that Iming arrived.

(3.74) bsiyaq ka k-n-pig-an phpah nii da
long.STAT.FIN NOM STAT.NFIN-PFV-wrinkled-GER flower.GEN PROX CS
'These flowers have been dead for a long time.' (lit. 'These flowers' being dead is long.')

[16] Ca- reduplication is also applied to the stem in Amis (Imanishi 2009:86, Liu 2011:31).

(3.75) l<m>ingis s<n>liq-an=mu rway-un=na
cry<AV> destroy<PFV>-GER=1SG.GEN adore-PAT.NMLZ=3SG.GEN
ka laqi nii
NOM child PROX
'This child is crying because of my breaking his toy.'

(3.76) ini=ku kla ka d<n>hq-an Iming
NEG=1SG.NOM know.AV.NFIN NOM arrive<PFV>-GER Iming.GEN
'I did not know about Iming's arrival (lit. arriving).'

(3.77) ini=ku kla dhuq ka Iming
NEG=1SG.NOM know.AV.NFIN arrive.AV NOM Iming
'I did not know that Iming had arrived.'

Telic verbs interact with the perfective aspect marker *<n>* in an intriguing way that atelic verbs do not. Namely, perfective gerund forms of telic verbs may refer to the event itself (3.78) or the resulting state of the said event (3.79). Which interpretation obtains is dependent on whether the action itself entails any temporal duration (i.e., accomplishment vs. achievement distinction), as well as external factors like the presence/absence of a time expression and functional items indicating perfectivity (e.g., perfective *wada*) and/or change of state (sentence-final particle *da*). The interaction between lexical aspect and grammatical aspect in gerunds is further explored in §3.4.1.

(3.78) dha idas
two months
p-n-hig-an=na sapah=na ka Iming
CAUS-PFV-stand-GER=3SG.GEN house=3SG.GEN NOM Iming
'It took Iming two months to build her house.' (lit. 'Iming's building her house is two months.')

(3.79) wada dha hnkawas
PFV two year
p-n-hig-an=na sapah ka Kuras nii da
CAUS-PFV-stand-GER=3SG.GEN house.OBL NOM Kuras PROX CS
'It has been two years since Kuras built the house.' (lit. 'Kuras' having built the house has been two years.')

3.3.2.2. Subjunctive mood

While indicative gerunds are used for known states and events that have already taken place or are certain to take place, subjunctive gerunds denote hypothetical events or states. Thus, they often co-occur with predicates of manipulation and speech act verbs, such as *ddug* 'urge', *skdug* 'instigate', *durum* 'request', *rngag* 'speak (to)', *dakar* 'forbid', and *qlahang* 'warn'. Other verbs accompanied by subjunctive gerunds, like *taga* 'wait' and *sdhug* 'arrange', also entail reference

to expectations and hypothetical events. Such events may or may not have already occurred, and may or may not take place in the future.

(3.80) d<m>udug knan g-geuguy=mu pila bubu=mu
urge<AV> 1SG.OBL SBJV.GER-steal=1SG.GEN money.OBL mother=1SG.GEN
ka Kuras
NOM Kuras
'Kuras urged me to steal my mother's money.' (lit. 'Kuras pushed me for my stealing my mother's money.')

(3.81) Utux Baraw o, m-dakar mnan
spirit above TOP AV-forbid 1PL.EXCL.OBL
t-tlung hiyaan
SBJV.GER-touch 3SG.OBL
'Lord forbade us from touching it (= the fruit of the tree of life)' (*Soyang Patas*, Genesis 3:3)

(3.82) nii=ku t<m>aga e-iyah=su
PROG.PROX=1SG.NOM wait<AV> SBJV.GER-wait=2SG.GEN
'I am waiting for you to come.' (lit. 'I am waiting for your coming.')

In the absence of such predicates, subjunctive gerunds usually denote intentions or purposes.

(3.83) k-k-brax hii=mu ka t<m>alang=ku
SBJV.GER-STAT.NFIN-strong body=1SG.GEN NOM run<AV>=1SG.NOM
'I run for my health.' (lit. '(The reason why) I run is my body's being strong.')

(3.84) M-iyah paah baraw
AV-come from above
q-qita=na alang ni daka s<n>alu seediq nii
SBJV.GER-see=3SG.GEN city and tower make<PFV.PV> person.GEN PROX
'He (= Lord) came down to see the city and the tower that these people built.' (lit. 'He came down for his seeing the city…') (*Soyang Patas,* Genesis 11.5)

3.3.3. Distribution

Both indicative and subjunctive gerunds appear in various syntactic slots including the nominative argument position (3.85), the oblique argument position (3.86), and the predicate position (3.87). Note that the gerund's position in each sentence is parallel to that of the second person free pronoun *isu* in (3.88–90). The form of the pronoun is *isu* (2SG) in (3.88) and (3.90), whereas it is *sunan* (2SG.OBL) in (3.89).

(3.85) tga-un=mu ka e-iyah=su
wait-PV=1SG.GEN NOM SBJV.GER-come=2SG.GEN
'I am waiting for you to come.' (lit. 'I am waiting for your coming.')

(3.86) nii=ku t<m>aga e-iyah=su
PROG.PROX=1SG.NOM wait<AV> SBJV.GER-come=2SG.GEN
'I am waiting for you to come.' (lit. 'I am waiting for your coming.')

(3.87) e-iyah=su ka nii=mu tga-un
SBJV.GER-come=2SG.GEN NOM PROG.PROX=1SG.GEN wait-PV
'It is for you to come that I am waiting.' (lit. 'What I am waiting for you is your coming.')

(3.88) tga-un=misu (ka isu)
wait-PV=1SG.GEN:2SG.NOM NOM 2SG
'I am waiting for you.'

(3.89) nii=ku t<m>aga sunan
PROG.PROX=1SG.NOM wait<AV> 2SG.OBL
'I am waiting for you.'

(3.90) isu ka nii=mu tga-un
2SG NOM PROG.PROX=1SG.GEN wait-PV
'It is you that I am waiting for.'

Gerunds also appear in default and postposed adjunct positions. The default adjunct position is post-verbal and before the *ka*-marked pivot (3.91a). Postposed adjuncts are sentence-final, following the pivot and a prosodic pause (3.91b). It remains to be seen whether postposing results in a slightly different semantic emphasis.

(3.91) a. l<m>ingis t<n>kur-an=na ka laqi
cry<AV> fall<PFV>-GER=3SG.GEN NOM child
b. l<m>ingis ka laqi, t<n>kur-an=na
cry<AV> NOM child fall<PFV>-GER=3SG.GEN
'The child$_i$ cried because he$_i$ fell.' (lit. 'The child cried because of his falling.')

Finally, gerunds can also be the object of a preposition like *asaw* 'because of'.

(3.92) asaw bi h<n>qil-an=na ka naqih kuxul=mu
because.of INT die<PFV>-GER=3SG.GEN NOM bad.STAT.FIN feelings=1SG.GEN
'It is because of his death (lit. his dying) that I am sad.'

To sum up, this section provided an overview of the morphology, semantics, and distribution of Truku gerunds. Indicative gerunds are formed with *<n>…-an* (with an added nonfinite stative marker *k-* for stative verbs) in the perfective aspect, and with *-un* in the future tense reading. Subjunctive gerunds are formed via Ce- reduplication. Semantically, a gerund refers to an event or a state resulting from an event, but not to propositions. Gerunds occur in a number of syntactic slots including the nominative pivot, oblique argument, predicate, adjunct, and object of a preposition.

3.4. Syntactic behavior of Truku gerunds

This section consists of a detailed description of the constructions under investigation, Truku gerunds. In §3.4.1, I will present the interplay of grammatical aspect marking and lexical aspect

in gerunds. It will be shown that, though perfective marking is commonly present in gerunds, it has different semantic effects depending on the verb's telicity. §3.4.2 will focus on case assignment patterns within gerunds. Specifically, no nominative case is assigned, while the subject (Actor) and the object (Patient) of the gerund are consistently marked genitive and oblique, respectively. Simultaneously, I will demonstrate that the subject of a gerund must be the Actor. Moreover, the subject cannot be omitted except in some marginal cases. I will argue that the Actor-subject requirement is a natural consequence of gerunds' lack of voice alternation. §3.4.3 first introduces the phenomenon of external possession. I will then demonstrate that it applies to gerunds in much the same ways as it does to other nominals. §3.4.4 explores negation strategies used in gerunds, which essentially align them with noun phrases rather than verb phrases. In contrast, §3.4.5 highlights verb-like features of gerunds, showing that the same manner modification strategies are employed for gerunds and prototypical verbal constructions. Finally, §3.4.6 investigates the co-occurrence possibilities of gerunds and a number of functional items. Elsewhere in the language's grammar, some of these items are exclusively found with nouns while others occur exclusively with verbs. Gerunds' varied behavior further confirms their in-between-category nature.

3.4.1. Grammatical and lexical aspect

In §3.3, we observed that Truku gerunds can be marked for perfective aspect. Furthermore, perfective aspect and the lexical aspect (a.k.a. Aktionsart) conspire to determine a gerund's semantic reference.

Vendler (1957, 1967) classifies English verbs into four categories according to their lexical aspect ("time schemata," in his terms): state, activity, accomplishment, and achievement (Table 3.2).[17] They differ in three aspects: (i) involvement of "change," (ii) telicity, and (iii) durativity.

Table 3.2. Four major verb categories

	change		no change
	telic	atelic	
durative	accomplishment	activity	stative
non-durative	achievement	---	

[17] Comrie (1976) adds the "semelfactive" category to this classification, defined as "a situation that takes place once and once only (e.g. one single cough)," as opposed to "iterative," which is "a situation that is repeated (e.g. a series of coughs)." I will not discuss semelfactive verbs in Truku since I do not have sufficient data.

Stative verbs can be fundamentally separated from non-stative verb categories in that the former require no change in situation and the latter do. Additionally, stative verbs contrast with non-stative ones in their interaction with grammatical aspect in many languages. For instance, only non-stative verbs can appear in progressive aspect in English. Thus, *Ben was reading a book* is grammatical while **Sarah is knowing the truth* is not. In some languages that grammatically code perfective vs. imperfective aspect, perfective forms of stative verbs can indicate the beginning of a state. Spanish verbs like *conocer* 'know' exemplify this point. With the perfective past form *conocí*, *conocí a Pedro hace muchos años* translates as 'I got to know Pedro many years ago' (Comrie 1976).

In addition to the stative vs. non-stative division, lexical aspect is defined in terms of two properties: telicity (i.e., presence/absence of a natural end point) and durativity (whether the action entails any duration of time, or is punctual). Activities such as *walking*, *running*, and *painting* and states such as *knowing*, *believing*, and *loving* have no natural end point, so that it is possible to say 'I walked *for an hour*' or 'I knew him *for several years*.' On the other hand, it is quite unnatural to combine such time-span adverbials with accomplishments (e.g., '*I ran a mile *for eight minutes*' and '*I walked to school *for half an hour*') or with achievements (e.g., '*I reached the summit *for two days*' and '*I won a game *for four hours*'). Conversely, time-frame adverbials can be combined with telic verbs ('I ran a mile *in eight minutes*' and 'I reached the summit *in two days*') but not with atelic ones ('*I ran *in an hour*'[18] and '*I knew him *in eight years*').

Accomplishment verbs diverge from achievement verbs in that they denote actions that are ongoing at any time for the duration of such time spans. For example, the accomplishment verb 'run (x distance)' presupposes duration of the activity. *It took me eight minutes to run a mile* entails that *I was running* at any time within those eight minutes. In contrast, achievement verbs are non-durative, and are not compatible with the progressive aspect. Therefore, if it took me two days to reach the summit, I cannot say that *I was reaching the summit* on day one. Activity verbs, like accomplishment verbs, are durative. Thus, with 'run' (with no phrases indicating a distance or a goal, which implies the existence of an end point), we can say that *I ran for an hour* means I was running at any point of time within that hour. In other words, suppose I have just

[18] Note that many verbs (e.g., 'run', 'walk', 'build', 'draw') have both activity and accomplishment usages.

finished running for an hour; I can say *I was running half an hour ago*. Note that stative verbs cannot occur in the progressive aspect, even though they entail durativity. That is, I could not say **I was knowing John four years ago,* even if I have known him for eight years (Comrie 1976:42).

In Truku, most state-denoting verbs have a unique set of morphology (*m-* in finite forms and *k-* in nonfinite forms) and usually do not inflect for voice. I have been referring to these verbs as stative verbs. However, there are some verbs that do not bear stative verb morphology, despite being semantically stative. For instance, *kla* 'know' and *(s)kuxul* 'like'[19] inflect for both AV (3.93a, 3.94a) and PV (3.93b, 3.94b).

(3.93) a. m-kla=ku sunan
AV-know=1SG.NOM 2SG.OBL
b. kla-un=misu
know-PV=1SG.GEN:2SG.NOM
'I know you.'

(3.94) a. s<m>kuxul uyas bubu=na ka Rubiq
like<AV> song.OBL mother=3SG.GEN NOM Rubiq
b. kuxul Rubiq ka uyas bubu=na
like<PV> Rubiq.GEN NOM song mother=3SG.GEN
'Rubiq likes her mother's songs.'

Due to insufficient data, I will not discuss the latter type of verb in relation to gerunds in Truku. Therefore, I will continue the practice of using the term "stative verbs" to refer to those state-denoting verbs which are also members of the morphologically defined lexical class.

The semantic effects of the perfective marker *<n>* on gerunds hinge on the lexical aspect of the verb. Stative verbs naturally form state-denoting gerunds.

(3.95) bsiyaq ka k-n-pig-an phpah nii da
long.STAT.FIN NOM STAT.NFIN-PFV-wrinkled-GER flower.GEN PROX CS
'These flowers have been dead for a long time.' (lit. 'These flowers' being dead is long.')

Perfective forms of achievement (telic and non-durative) verbs may result in state-denoting gerunds, such that *d<n>hq-an* (< *dhuq* 'arrive') refers to the state of 'having arrived', and *h<n>qil-an* (< *huqil* 'die') to the state of 'having died'. Other achievement verbs include *hlay* 'find', *alax* 'quit, give up', and *qrap* 'capture'.[20] Expressions of time duration associated with

[19] *(S)kuxul* 'like' is an irregular verb whose form is zero-marked in its PV form. In addition, the initial consonant /s/ is present in the AV form *s<m>kuxul* but is lacking in the PV form *kuxul*.

[20] These verbs are classified as achievement verbs according to their interaction with the perfective marker. However, translated equivalents of some Truku verbs may be classified differently in English. Note that such classification is language-specific to a certain degree.

telic verb gerunds unequivocally point to the duration of such a state that emerges *after* the occurrence of a certain event (e.g., the state of Kuras' having arrived is a result of his arrival). For instance, the only interpretation available to (3.96) is 'it has been three days since Kuras arrived'; it cannot mean 'it took three days for Kuras to arrive'. Note that in Truku, time expressions are formally indistinguishable between being of the time-frame and the time-span types. Duration is frequently expressed via a predicate, with the event/state in the form of a gerund in the nominative argument position[21].

(3.96) mk-tru jiyax ka d<n>hq-an Kuras
be.x.days-three day NOM arrive<PFV>-GER Kuras.GEN
'It has been three days since Kuras arrived.' (lit. 'Kuras' having arrived is three days.')

(3.97) mk-tru jiyax ka q<n>rap-an pais knan
be.x.days-three day NOM capture<PFV>-GER enemy.GEN 1SG.OBL
'I have been/was in the enemy's captivity for three days.' (lit. 'The enemy's having caught me is three days.')

Yet, in the absence of time expressions, perfective gerunds of achievement verbs are ambiguous as to whether they refer to the event itself or the state resulting from an event. 'Arrival' in (3.98) can either mean 'the fact that Iming arrived (i.e., Iming's having arrived)' or 'the circumstances under which Iming arrived (i.e., Iming's arrival)'.

(3.98) ini=ku kla d<n>hq-an Iming
NEG=1SG.NOM know.AV.NFIN arrive<PFV>-GER Iming.GEN
'I did not know about Iming's arrival (lit. arriving).'

Activity (atelic and durative) verbs contrast with achievement verbs in that the time expressions associated with them simply refer to the duration of the event itself. There are numerous activity verbs, such as *ksa* 'walk', *talang* 'run', and *sais* 'sew'.

(3.99) shiga o, rima tuki ka s<n>qit-an=mu qhuni
yesterday TOP five hour NOM cut<PFV>-GER=3SG.GEN tree.OBL
'Yesterday, I cut trees for five hours.' (lit. 'Yesterday, my cutting trees is five hours.')

In contrast, perfective-marked gerunds of accomplishment (telic and durative) verbs can be state- or event-denoting depending on the context, even with a time adverbial. The gerund *p-n-hig-an* 'building' in (3.100) denotes the event of the subject's (i.e., Iming's) building the house,

[21] There are instances like (3.100–101) where the gerund seems to lack the nominative case-marker. I argue that this is due to the co-occuring external possession, which introduces an additional *ka*-marked phrase (see §3.4.3 and §3.7). Native speakers have a strong preference to have no more than one overt *ka* per clause, often deleting others.

while the phonologically identical gerund in (3.101) denotes the state of the subject (i.e., Kuras) having built the house. Thus, the time-frame vs. time-span interpretation of the associated time expression also varies depending on the context. It is the perfective marker *wada* and the change of state particle *da* that serve to disambiguate between the two usages. These functional items in (3.101) indicate that the time span since the completion of the event up until the present time *has been* two years. On the other hand, the time frame for the event in (3.100) will not change over time; it will always remain two months.

(3.100) dha idas
two months
p-n-hig-an=na sapah=na ka Iming
CAUS-PFV-stand-GER=3SG.GEN house=3SG.GEN NOM Iming
'It took Iming two months to build her house.' (lit. 'Iming's building her house is two months.')

(3.101) **wada** dha hnkawas
PFV two year
p-n-hig-an=na sapah ka Kuras nii **da**
CAUS-PFV-stand-GER=3SG.GEN house.OBL NOM Kuras PROX CS
'It has been two years since Kuras built the house.' (lit. Kuras' having built ths house has been two years.')

The use of *wada* and *da* in combination with an achievement verb does not affect the interpretation of time expressions: they have the time-span reading with them (3.102) or without them (3.103). Whereas *wada* and *da* in (3.102) introduce the sense of change of state (in that five months have passed), the state of 'my having quit smoking for three months in the past' in (3.103) is time-stable.

(3.102) **wada** rima idas ka n-lax-an lumak Kuras **da**
PFV five month NOM PFV-quit-GER smoke.OBL Kuras.GEN CS
'It has been five months since Kuras quit smoking.' (lit. 'Kuras' having quit smoking has been five months.')

(3.103) tru idas ka n-lax-an=mu m-kan lumak han,
three month NOM PFV-quit-GER=1SG.GEN AV-eat smoke.OBL a.while
m-kan=ku duri ka sayang
AV-eat=1SG.NOM too NOM now
'I quit smoking for three months; now I smoke again.' (lit. 'My having quit smoking was three months…')

To summarize, gerunds formed on non-stative verbs can refer to the event itself or the state of an event having been completed. However, the state reading is available to accomplishment and achievement verbs only.

Table 3.3. Availability of event and state readings according to verbal lexical aspect

	stative	non-stative		
		telic	atelic	
		activity	accomplishment	achievement
event reading	no	yes	yes	yes
state reading	yes	no	yes	yes

3.4.2. Case assignment and the lack of voice alternation

Recall that a variety of thematic roles can serve as the syntactic pivot. This is what makes the Philippine-type voice systems unique. As discussed in Chapter 2, in the canonical verbal constructions of Truku, nominative case is regularly assigned to the argument whose thematic role corresponds to the voice morphology. In NAV sentences, the Actor is marked for genitive case. Crucially, no other participant but the Actor of a transitive verb receives the genitive case. As discussed previously, genitive case on NPs is zero-marked. Furthermore, the nominative and genitive pronoun pairs can be distinguished only in the first person singular, third person singular, and third person plural. In (3.104) with an AV verb, the Actor appears in the nominative form *=ku*, while in (3.105) with a PV verb, it appears in the genitive form, *=mu*.

(3.104) s<m>ipaq=ku sunan (Actor Voice)
hit<AV>=1SG.NOM 2SG.OBL
'I hit you.'

(3.105) bhang-an=mu ka uyas qbhni (Patient Voice)
hear-PV=1SG.GEN NOM song bird
'I hear the bird song.'

In gerunds, genitive case is assigned to the Actor of not only transitive verbs (3.106) but also of unergative verbs (3.107) and unaccusative verbs (3.108). I will refer to these participants as the *subject* of the gerund.

(3.106) mk-tru jiyax ka q<n>rp-an=dha knan
be.x.days-three day NOM catch<PFV>-GER=3PL.GEN 1SG.OBL
'It has been three days since they caught me.' (lit. 'Their having caught me is three days.')

(3.107) kingal iyax tuki ka n-ksa-an=mu
one interval hour NOM PFV-walk-GER=1SG.GEN
'I walked for an hour.' (lit. 'My walking is an hour.')

(3.108) l<m>ingis t<n>kur-an=na ka laqi
cry<AV> fall<PFV>-GER=3SG.GEN NOM child
'The child$_i$ cried because he$_i$ fell.' (lit. 'The child cried because of his falling.')

The Patient of a transitive gerund bears oblique case (3.109). In contrast, in verbal constructions, the Patient bears oblique case in AV sentences (3.110), but it must be nominative in PV sentences (3.111).

(3.109) dha bi tuki ka tga-un=mu sunan
two INT hour NOM wait-GER=1SG.GEN 2SG.OBL
'I will wait for you for (no more than) two hours.' (lit. 'My waiting for you is two hours.')

(3.110) nii=ku t<m>aga sunan
PROG.PROX=1SG.NOM wait<AV> 2SG.OBL
'I am waiting for you.'

(3.111) a. tga-un=misu (ka isu)
wait-PV=1SG.GEN:2SG.NOM NOM 2SG
b. *tga-un=mu sunan
wait-PV=1SG.GEN 2SG.OBL
'I will wait for you.'

In addition, the Patient cannot be marked genitive in gerunds. In other words, they cannot be the subject. Thus, I will refer to the Patient as the *object* of a gerund.

(3.112) *mk-tru jiyax ka q<n>rp-an=mu
be.x.days-three day NOM catch<PFV>-GER=1SG.GEN
IM: 'It has been three days since I was caught.' (lit. 'My having been caught is three days.')

It should also be noted that, in principle, no argument of a gerund can be marked as nominative. External possession presents an exception to this. Briefly, external possession is a phenomenon in which the possessor alone, rather than an entire possessive phrase, appears in the nominative NP position. Meanwhile, the possessum surfaces as an adjunct to the predicate. When external possession is applied to gerunds, the subject is treated on par with the possessor, occurring with the nominative marker *ka* and severed from the rest of the gerund. Nominative marking in gerunds will be discussed in detail in §3.4.3.

Not only is the subject of a gerund restricted to the Actor, but it is also mandatory for gerunds to have an overt subject. When the identity of the subject is unknown, a dummy subject (either third person singular or plural) is inserted.

(3.113) m-saang h<n>muk-an*(=dha) knan
STAT.FIN-angry lock.up<PFV>-GER=3PL.GEN 1SG.OBL
ka bubu=mu
NOM mother=1SG.GEN
'My mother is angry because they imprisoned me.' (lit. My mother is angry because of their locking me up.')

Usually, the subject must be overtly expressed even when coreference is retrievable from the context.

(3.114) l<m>ingis t<n>kur-an*(=na) ka laqi
cry<AV> fall<PFV>-GER=3SG.GEN NOM child
'The child$_i$ cried because he$_i$ fell.'[22] (lit. 'The child cried because of his falling.')

Nonetheless, there are occasional cases where the subject can be omitted. This is especially the case with subjunctive gerunds that co-occur with a verb of causation, speech act, or expectation (e.g., *durum* 'request' in 3.115), where the subject ('you') has already been mentioned.

(3.115) d<m>urum=ku sunan e-adas(=su) Kuras
request<AV>=1SG.NOM 2SG.OBL SBJV.GER-bring=2SG.GEN Kuras.OBL
'I asked you to bring Kuras.' (lit. 'I asked you for your bringing Kuras.')

The fact that the subject can only be the Actor means that the gerund lacks voice alternation. Compare this situation with English, which allows gerunds to be passivized, so that the Patient serves as the subject (3.116).

(3.116) The President being kidnapped shocked the entire nation.

The important claim that I make here is that, *even though gerund markers are homophonous with voice markers, they do not have voice-altering functions*. If they did, we would expect the Patient to be marked as nominative and not oblique.

3.4.3. Nominative marking in gerunds

As demonstrated in §3.4.2, the subject (i.e., Actor) of gerunds is normally marked for genitive case (3.106 and 3.108, repeated as (3.117) and (3.118) here).

(3.117) mk-tru jiyax ka q<n>rp-an=dha knan
be.x.days-three day NOM catch<PFV>-GER=3PL.GEN 1SG.OBL
'It has been three days since they caught me.' (lit. 'Their having caught me is three days.')

[22] Although coreference between the subject of the gerund and the sentential pivot is not required, they are understood to be identical in this context; the only logical interpretation is that the person that fell is also the one crying.

(3.118) l<m>ingis t<n>kur-an=na ka laqi
cry<AV> fall<PFV>-GER=3SG.GEN NOM child
'The child$_i$ cried because he$_i$ fell.' (lit. 'The child cried because of his falling.')

In this section, I present exceptions to this pattern where the subject is marked nominative. In such instances, the gerundive verb bears an obligatory genitive clitic that matches the subject in person and number. This pattern is available to both indicative (3.119) and subjunctive (3.120) gerunds alike. In each of the examples below, the (a)-sentence represents a gerund with a genitive-marked subject, and the (b)-sentence represents one with a nominative-marked subject.

(3.119) a. tru jiyax ka q<n>rap-an knan pais
three day NOM capture<PFV>-GER=3PL.GEN 1SG.OBL enemy.GEN
b. tru jiyax q<n>rap-an*(=dha) knan ka pais
three day capture<PFV>-GER=3PL.GEN 1SG.OBL NOM enemy
'It has been three days since the enemies caught me.' (lit. 'The enemies' having caught me is three days.')

(3.120) a. gisu=ku t<m>aga
PROG.PROX=1SG.NOM wait<AV>
e-iyah sapah Walis
SBJV.GER-come house.OBL Walis.GEN
b. gisu=ku t<m>aga
PROG.PROX=1SG.NOM wait<AV>
e-iyah*(=na) sapah ka Walis
SBJV.GER-come=3SG.GEN house.OBL NOM Walis
'I am waiting for Walis to come home.' (lit. 'I am waiting for Walis' coming home.')

In both (3.119b) and (3.120b), the subject (*pais* 'enemies' and *Walis*, respectively) are preceded by the nominative case marker *ka*. *Pais* in (3.119b) is co-indexed on the verb 'capture(-ing)' as the third person plural genitive clitic *=dha*. Likewise, *Walis* is coreferential with the genitive clitic *=na* found on the verb 'come(-ing)'.

For instances such as (3.119b), one may be tempted to treat *pais* 'enemy' as the syntactic pivot of the sentence that receives nominative case from the matrix predicate. On the contrary, I assume the nominative-marked subject to be internal to the gerund. This is because nominative-marked subjects of gerunds lack typical properties of a true sentential pivot.

First, nominative marking on the gerundive subject occurs even with matrix clauses that already have a nominative-marked pivot, provided that it would not result in two *ka*-marked phrases. For example in (3.120b), the pivot of the matrix clause appears as the first person nominative clitic *=ku*. Recall from §2.4 that time expressions behave in a similar manner (3.121–

122). They may occur sentence-finally with the *ka* marker in the presence of a cliticized or topicalized pivot. While the latter is semantically and syntactically the subject of the clause, the "nominative" marking on the time expression is merely morphological, not a reflection of its status as a syntactic pivot.

(3.121) m-iing=ku Kuras ka sayang
AV-search=1SG.NOM Kuras.OBL NOM now
'I am looking for Kuras right now.'

(3.122) Hana o, sprang aji me-tabug huling
Hana TOP intentionally NEG AV.IRR-feed dog.OBL
ka saman
NOM tomorrow
'Hana will intentionally not feed the dog tomorrow (e.g., in order to punish it).'

Second, no element belonging to the matrix clause may intervene between nominative-marked subjects of gerunds and the gerundive verb (in addition to all of its complements and adjuncts). Therefore, temporal expressions like *suni* 'a short while ago' and *shiga* 'yesterday' that appear in this position are necessarily interpreted as modifying the gerund and not the matrix predicate. Compare:

(3.123) bhang-an=mu d<n>hq-an=na suni ka Iming
hear-PV=1SG.GEN arrive<PFV>-GER=3SG.GEN short.while.ago NOM Iming
'I hear that Iming arrived a short while ago.' (lit. 'I hear about Iming's having arrived a short while ago.')
NOT 'I heard a short while ago that Iming arrived.'

(3.124) bhang-an=mu suni d<n>hq-an=na ka Iming
hear-PV=1SG.GEN short.while.ago arrive<PFV>-GER=3SG.GEN NOM Iming
'I heard a short while ago that Iming arrived.' (lit. 'I heard a short while ago about Iming's having arrived.')

Note that the word order of time expressions is usually relatively free.

(3.125) a. puy-an sari laqi **kdjiyax** ka hini
cook-LV taro.OBL child.GEN every.day NOM here
b. puy-an **kdjiyax** sari laqi ka hini
cook-LV every.day taro.OBL child.GEN NOM here
c. puy-an sari **kdjiyax** laqi ka hini
cook-LV taro.OBL every.day child.GEN NOM here
'The child cooks taro here every day.'

Despite appearing separated from each other, the nominative-marked subject and the rest of the gerund somehow need to be adjacent to each other.

One may ask, then, "what is the pivot of the entire sentence in (3.123–124), if not *Iming*?" I argue that it is the whole of the gerund 'Iming's having arrived a short while ago', even though it lacks nominative case-marking. Note that the nominative case-marker borne by the gerund is nearly always lost when its subject takes on *ka*. Compare (3.124) with (3.126) below, the latter of which has the subject of the gerund in genitive case.

(3.126) bhang-an=mu suni ka d<n>hq-an Iming
hear-PV=1SG.GEN short.while.ago NOM arrive<PFV>-GER Iming.GEN
'I heard a short while ago that Iming arrived.' ' (lit. 'I heard a short while ago about Iming's having arrived.')

This is not a trait particular to these constructions; rather, it is due to the general preference for a single overt *ka* per clause. Speakers may accept having the additional nominative marker in these sentences, but they typically express feeling as though they contain "too many *ka*'s." As a point of comparison, refer to the set of examples in (3.127–128), which do not involve gerunds. The pivot of (3.128) is the statement in (3.127), which already contains one *ka* marking the Patient of the PV verb, *qhuni nii* 'these trees' as nominative. When this statement itself serves as clausal pivot of a sentence, as in (3.128a), it, too, can be preceded by *ka*. In this case, the original *ka* associated with *qhuni nii* is deleted. Alternatively, this noun phrase retains its case-marker, but the clausal pivot does not receive one (3.128b). It is highly unnatural for both to be marked with *ka* (3.128c).

(3.127) s<n>qit-an=mu ka qhuni nii
cut<PFV>-PV=1SG.GEN NOM tree PROX
'I cut these trees.'

(3.128) a. tru iyax tuki **ka** s<n>qit-an=mu **Ø** qhuni nii
b. tru iyax tuki **Ø** s<n>qit-an=mu **ka** qhuni nii
c. ?tru iyax tuki **ka** s<n>qit-an=mu **ka** qhuni nii
three interval hour NOM cut<PFV>-PV=1SG.GEN NOM tree PROX
'I have been cutting these trees for three hours.'

In turn, the gerund in (3.120a–b, repeated here as 3.129a–b) bears oblique case. One may not insert *ka* before the gerund (3.129c), suggesting that the matrix clause cannot assign nominative case to it.

(3.129) a. gisu=ku t<m>aga
PROG.PROX=1SG.NOM wait<AV>
e-iyah sapah Walis
SBJV.GER-come house.OBL Walis.GEN
b. gisu=ku t<m>aga
PROG.PROX=1SG.NOM wait<AV>
e-iyah*(=na) sapah ka Walis
SBJV.GER-come=3SG.GEN house.OBL NOM Walis
c. *gisu=ku t<m>aga
PROG.PROX=1SG.NOM wait<AV>
ka e-iyah sapah Walis
NOM SBJV.GER-come house.OBL Walis.GEN
'I am waiting for Walis to come home.' (lit. 'I am waiting for Walis' coming home.')

I take this phenomenon of nominative case-marking on gerundive subjects as cases of external possession in which the possessor and the possessum do not form a constituent. In Truku external possessor constructions (EPCs), the possessor, rather than the entire possessive phrase, appears nominative-marked. The possessum appears preceding *ka* at the right edge of the predicate. An example of an EPC is provided in (3.130a) along with its non-external possession counterpart (3.130b).

(3.130) a. wada=mu paq-un tunux=na ka Iming
PFV=1SG.NOM hit-PV head=3SG.GEN NOM Iming
b. wada=mu paq-un ka tunux Iming
PFV=1SG.NOM hit-PV NOM head Iming.GEN
'I hit Iming's head.'

When EPCs apply to gerunds, the subject is treated on par with the possessor, and all other components of the gerund as the possessum. The "possessum" minimally consists of the gerundive verb but also includes the Patient, as well as any modifiers and adverbs if present. This parallelism with possessive phrases is yet another aspect in which gerunds display nominal properties. I refer interested readers to the detailed illustration of EPCs provided in §3.7.

3.4.4. Negation strategies

This section investigates how negative gerunds are formed. Two primary negative markers, *ini* and *aji*, were introduced in §2.8. Both occur preverbally and host pronominal clitics. By virtue of being an auxiliary, *ini* requires that the following verb take nonfinite voice morphology. Being a preverb, *aji* takes finite verb forms. Moreover, *aji* is primarily used to refer to future events. *Ini* is usually used to negate states or events in the present time or the past. The key distinction between the two is that the former negates verb phrases, and the latter negates

propositions. Highly relevant to the current discussion of gerunds is the fact that *ini* does not negate nominal predicates (3.133–134).

(3.131) aji=mu tbug-un ka huling
NEG=1SG.GEN nurture-PV NOM dog
'I will not feed the dog.'

(3.132) ini=mu qdal-i n-hapuy shiga ka huling
NEG=1SG.GEN feed-PV.NFIN RES.NMLZ-cook yesterday NOM dog
'I did not feed the dog yesterday.'

(3.133) aji=ku kuyuh mp-hmut (ka yaku)
NEG=1SG.NOM woman AV.IRR-do.as.one.pleases NOM 1SG
'I am not an immoral woman.'

(3.134) *ini=ku kuyuh mp-hmut (ka yaku)
NEG=1SG.NOM woman AV.IRR-do.as.one.pleases NOM 1SG
IM: 'I am not an immoral woman.'

Negative indicative gerunds are formed with *aji* (3.137) and seem to occur only in the *-un* suffixed form.

(3.135) rima jiyax ka aji=ku me-talang
five day NOM NEG=1SG.NOM AV.IRR-run
'I will not run for five days.'

(3.136) rima jiyax ka ini=ku talang da
five day NOM NEG=1SG.NOM run.AV.NFIN CS
'I have not run in five days.'

(3.137) rima jiyax ka aji=mu tlam-un
five day NOM NEG=1SG.GEN run-GER
'I will not run for five days.' (lit. 'My not running is five days.')

(3.138) *rima jiyax ka aji=mu t<n>lam-an
five day NOM NEG=1SG.GEN run<PFV>-GER
IM: 'I have not run in five days.'

On the other hand, gerunds cannot contain *ini*.

(3.139) *rima jiyax ka ini=mu t<n>lam-an
five day NOM NEG=1SG.GEN run<PFV>-GER
IM: 'I have not run in five days.'

Like indicative gerunds, subjunctive gerunds may also contain the negative preverb *aji*. No instance of *ini* inside a subjunctive gerund has been found in my corpus.

(3.140) m-ekan=ku malu uq-un o,
AV-eat=1SG.NOM good.STAT.FIN eat-PAT.NMLZ TOP
aji=mu k-k-narux
NEG=1SG.GEN SBJV.GER-STAT.NFIN-sick
'I eat well so that I will not get sick.' (lit. 'I eat good food; it's for my not being sick.')

(3.141) rngag-an=mu aji=na g-gsuqi ka laqi=mu
speak-PV=1SG.GEN NEG=3SG.GEN SBJV.GER-late NOM child=1SG.GEN
'I told my child not to be late.' (lit. 'I spoke to my child about his not being late.')

To summarize, negative gerunds are formed by *aji* but not *ini*. However, *aji* is only compatible with non-perfective indicative and subjunctive morphology. The language appears to lack a means of forming negative gerunds in the perfective form, creating a paradigm gap.

3.4.5. Modification strategies

As we observed in §3.2, one of the key differences between derived nouns and gerunds in English is how they are modified. Specifically, derived nouns take adjectives (3.142b), while gerunds take adverbs (3.142c).

(3.142) a. John criticized the book **harshly**.
b. John's **harsh**/*harshly criticism of the book
c. John's **harshly**/*harsh criticizing the book

This section will demonstrate that the modification strategy used for Truku gerunds is one that is employed in verb phrases rather than noun phrases.

Truku lacks both the category of adjectives and productive adverb-deriving morphology. Underived nouns are modified using juxtaposition of state-denoting verbs[23]. Verbs usually follow the noun being modified, but the reverse order is also possible. A few stative verbs, including *malu* 'good', *naqih* 'bad', and *paru* 'big', appear more frequently in the pre-nominal slot than the post-nominal one.

(3.143) a. lukus [m-ruciq]
clothes STAT.FIN-dirty
b. [m-ruciq] lukus
STAT.FIN-dirty clothes
'dirty clothes'

[23] These modifying verb forms can be considered relative clauses.

(3.144) a. [malu] seediq
good.STAT.FIN person
b. seediq [malu]
person good.STAT.FIN
'good person'

Divergently, manner modification is achieved primarily through verb serialization. In manner serial verb constructions (SVCs)[24], the modifying verb regularly serves as V1, hosting pronominal clitics and carrying TAM morphology, as well as a voice marker corresponding to the thematic role of the sentential pivot. V2 is an action verb in its AV form. The relative ordering of the manner-denoting and action-denoting verbs cannot be reversed. In (3.145) and (3.146), the V1 *drmut* 'diligent' and *biyax* 'strong' modify the V2 *karag* 'clean' and *sipaq* 'hit', respectively.

(3.145) k-n-drmut-an=mu k<m>arag ka sapah=mu
STAT.NFIN-PFV-diligent-PV=1SG.GEN clean<AV> NOM house=1SG.GEN
'I cleaned my house diligently.'

(3.146) k-n-biyax-an=mu s<m>ipaq ka huling
STAT.NFIN-PFV-strong-PV=1SG.GEN hit<AV> NOM dog
'I hit the dog hard.'

Gerunds do not behave like prototypical nouns in terms of modification. In (3.147) and (3.148), the (a)-sentence is a simple gerund, and the (b)-sentence represents a failure to modify it via juxtaposition of a manner-denoting verb ('diligent' and 'strong', respectively).

(3.147) a. bhgay bi k<n>rag-an=mu
white.STAT.FIN INT STAT.NFIN-PFV-clean-GER=1SG.GEN
ka sapah=mu
NOM house=1SG.GEN
'My house is clean because of my cleaning.'

b. *bhgay bi k<n>rag-an=mu m-drmut
white.STAT.FIN INT clean<PFV>-GER=1SG.GEN STAT.FIN-diligent
ka sapah=mu
NOM house=1SG.GEN
IM: 'My house is clean because of my diligent cleaning.'

(3.148) a. m-huqil n-paq-an=mu ka huling
AV-die PFV-hit-GER=1SG.GEN NOM dog
'The dog died because of my hitting (it).'

b. *m-huqil n-paq-an=mu m-biyax ka huling
AV-die PFV-hit-GER=1SG.GEN STAT.FIN-strong NOM dog
IM: 'The dog died because of my hitting it hard (lit. my strong hitting).'

[24] See §2.8 for a general discussion of SVCs.

Just as in verbal constructions, manner modification for gerunds is achieved via verb serialization. Gerundive SVCs follow general patterns of verb serialization. A manner-denoting verb precedes the action-denoting verb being modified, and hosts the gerund marker, as well as the perfective marker if in the perfective aspect. The action-denoting V2 bears default finite AV morphology with no TAM or gerund marking.

(3.149) bhgay bi k-n-drmut-an=mu k<m>arag
white.STAT.FIN INT STAT.NFIN-PFV-diligent-GER=1SG.GEN clean<AV>
ka sapah=mu
NOM house=1SG.GEN
'My house is clean because I cleaned it diligently.' (lit. 'My house is clean because of my diligent-ing clean.')

(3.150) m-huqil k-n-biyax-an=mu s<m>ipaq ka huling
AV-die STAT.NFIN-PFV-strong-GER=1SG.GEN hit<AV> NOM dog
'The dog died because I hit it hard.' (lit. 'The dog died because of my hard-ing hit.')

3.4.6. Compatibility with functional items

Finally, gerunds' compatibility with some functional items will be assessed in this section. Some of these items co-occur with nouns and others with verbs, but typically not with both. Though this list of functional items is not complete, it serves to further elucidate the mixed nature of gerunds, characterized by both nominal and verbal properties.

3.4.6.1. Demonstratives *nii* and *gaga*

Demonstratives *nii* (proximal) and *gaga* (distal) are placed after noun phrases.

(3.151) supaw nii / gaga
apple PROX DIST
'this / that apple'

Demonstratives seem to be largely unable to modify gerunds (3.152–153). Instead, time adverbials are the preferred method of referring to a specific instance of an event (3.154–155).

(3.152) *malu bi p<n>hpah-an=na nii ka buxi=mu[25]
good.STAT.FIN INT bloom<PFV>-GER=3SG.GEN PROX NOM lily=1SG.GEN
IM: 'This blooming of my lily is very good (pretty).'

[25] Note that gerunds in (3.152–155) have nominative-marked subjects. These examples were chosen because genitive marking on the subject would result in a word order where the demonstrative is taken to modify the subject rather than the gerund. Compare (3.152) with (f.1):

(f.1) malu bi ka p<n>hpah-an [buxi=mu nii]
good.STAT.FIN INT NOM bloom<PFV>-GER lily=1SG.GEN PROX
'This lily of mine bloomed prettily.' (lit. 'The blooming of this lily of mine is very good.')

(3.153) *m-shjil k-nrx-an=na gaga ka Kuras
STAT.FIN-heavy STAT.NFIN-sick-GER=3SG.GEN DIST NOM Kuras
IM: 'That sickness (lit. being sick) of Kuras' was serious.'

(3.154) malu bi p<n>hpah-an=na sayang ka buxi=mu
good.STAT.FIN INT bloom<PFV>-GER=3SG.GEN now NOM lily=1SG.GEN
'This time my lily bloomed very prettily.' (lit. 'My lily's blooming this time is good.')

(3.155) m-shjil k-nrx-an=na snduray ka Kuras
STAT.FIN-heavy STAT.NFIN-sick-GER=3SG.GEN previous.time NOM Kuras
'Kuras' sickness that time was serious.' (lit. 'Kuras' being sick that time is serious.')

Nonetheless, there are a few exceptions. At this time, it is unclear what factors affect the acceptability of demonstratives with gerunds.

(3.156) ?q<n>driq-an=su gaga o,
escape<PFV>-GER=2SG.GEN DIST TOP
m-skluwi=nami balay
STAT.FIN-surprised=1PL.EXCL.NOM INT
'That escape of yours really surprised us.' (lit. 'That escaping of yours, we were really surprised.')

(3.157) g<n>sqi-an=su nii han o,
late<PFV>-GER=2SG.GEN PROX PART TOP
aji me-seeliq p<n>tas-an=su
NEG AV.IRR-ruin study<PFV>-RES.NMLZ=2SG.GEN
'Your tardiness this time (lit. this being late of yours) will not affect your grade.'

3.4.6.2. *Asaw* 'because of / for'

The preposition *asaw* 'because of / for' normally takes a noun phrase as its complement[26]. *Asaw* phrases can be a predicate (3.158) or an adjunct (3.159). In the latter case, it is normally in the preverbal position.

(3.158) asaw bubu=na ka s-lingis=na
because.of mother=3SG.GEN NOM CV-cry=3SG.GEN
'The reason he cries is because of his mother.'

(3.159) asaw yami s<m>alu dha sapah ka laqi=mu
for 1PL.EXCL make<AV> two house.OBL NOM child=1SG.GEN
'My child built two houses for us.'

[26] An elderly informant claims that *asaw* should only be used in situations where a participant is negatively affected. However, it is also frequently used with a benefactive connotation, thus translated as 'for'.

In contrast, *asaw* may not be followed by a verb phrase.

(3.160) *m-ruciq ka lukus=mu o,
STAT.FIN-dirty NOM clothes=1SG.GEN TOP
asaw bi s-m-ruciq=su
because.of INT CAUS-STAT.FIN-dirty=2SG.NOM
IM: 'My clothes are dirty, it is because you made them dirty.'

(3.161) *asaw bi m-huqil ka hiya
because.of INT AV-die NOM 3SG
naqih kuxul=mu
bad.STAT.FIN feelings=1SG.GEN
IM: 'Because he died, I am sad.'

Gerunds resemble noun phrases in terms of their ability to be the complement of *asaw*.

(3.162) m-ruciq ka lukus=mu o,
STAT.FIN-dirty NOM clothes=1SG.GEN TOP
asaw bi s-n-rciq-an=su
because.of INT CAUS-PFV-dirty-GER=2SG.GEN
'My clothes are dirty, it is because of your making (them) dirty.'

(3.163) asaw bi h<n>qil-an=na
because.of INT die<PFV>-GER=3SG.GEN
ka naqih kuxul=mu
NOM bad.STAT.FIN feelings=1SG.GEN
'It is because of his death (lit. dying) that I am sad.'

3.4.6.3. *Siida* 'when'

The function of *siida* is to create a temporal adverbial clause. It occurs either following a finite clause (3.164) or within a finite clause between the verb phrase and the nominative NP (3.165). *Siida*-clauses specify the time at which the event denoted by the main clause takes place.

(3.164) rudan bi ka Aborahamo siida…
elderly INT NOM Abraham when
'When Abraham was old…' (*Soyang Patas,* Genesis 21:2)

(3.165) saw mp-k-stuq siida ka tama=mu o,…
SIM IRR-STAT.NFIN-severed when NOM father=1SG.GEN TOP
'When my father was about to die…' (*Soyang Patas*, Genesis 50:5)

Unlike typical verb phrases (3.166) or finite clauses, gerunds may not be followed by *siida* (3.167).

(3.166) qduriq=su siida o,
escape.AV=2SG.NOM when TOP
m-n-kluwi=nami balay
STAT.FIN-PFV-surprised=1PL.EXCL.NOM INT
'When you escaped, we were very surprised.'

(3.167) *q<n>driq-an=su siida o,
escape<PFV>-GER=2SG.GEN when TOP
m-n-kluwi=nami balay
STAT.FIN-PFV-surprised=1PL.EXCL.NOM INT
IM: ‘At the time of your escaping, we were very surprised.’

3.5. Interim summary

So far, this chapter has described morphological and syntactic properties of Truku gerunds. To recapitulate, there are two sets of gerunds with distinctive morphology and semantics: indicative and subjunctive. Indicative gerunds refer to events that have either taken place, or those that the speaker is certain will take place. They are marked with the affixes *-un* (and occasionally *-an*) in the non-perfective aspect, and with *<n>…-an* in the perfective aspect. These markers are homophonous with voice markers introduced in Chapter 2: the PV/perfective marker *<n>*, the PV non-perfective marker *-un*, and *-an*, which is used to mark both PV and LV.

On the other hand, subjunctive gerunds refer to hypothetical events. They are formed by Ce-reduplication of the verb stem. Indicative and subjunctive gerunds have largely overlapping distributional patterns: nominative argument position, the oblique argument position, the predicate position, the adjunct positions, and the object of prepositions. Thus, gerunds resemble prototypical nouns in their distributional patterns.

In §3.4, I demonstrated that Truku gerunds are characterized by an array of syntactic behavior that is both noun-like and verb- or sentence-like. The subject is generally marked for genitive case, much as the possessor of a noun is. Like noun phrases, gerunds mostly do not contain nominative-case marked elements. Optional nominative case-marking on the subject of a gerund is analyzed as external possession, where the gerund is treated similarly to a possessive (noun) phrase. Of the two major negative markers in the language, *aji* and *ini*, only the former can negate nouns and gerunds.

At the same time, indicative gerunds often show both aspectual marking and interaction with lexical aspect. Like a sentence, they require the subject, and the object is marked for oblique case. While nouns are modified via juxtaposition of a stative verb, a gerund is modified via verb serialization. Despite being unable to take demonstratives, gerunds can co-occur with the preposition *asaw* ‘because of/for’, which typically takes a noun phrase as its complement. Their lack of sententiality is also found in their inability to combine with *siida* ‘when’ and form temporal clauses.

3.6. Other types of nominalization

This section is an overview of various types of nominalization in Truku besides gerunds. Though the list is not exhaustive, it is intended to frame gerunds within other instances of nominalization both morphologically and syntactically. It will be apparent that these subcategories manifest different levels of productivity.

Similar examples have generally been treated as "lexical nominalization" in Formosan literature. However, for reasons stated in §3.2, I will not categorically treat them as such. First, there is a terminological ambiguity; lexical nominalization may involve items listed in the lexicon in their affixed forms and independent of their base verbs. Alternatively, it could refer to a productive morphological process that derives new items within the lexicon. Furthermore, some of the nominalizations below show ambiguous status between established and accepted lexical items, on the one hand, and verbal or gerundive forms created with productive morphology, on the other. Specifically, many of the nominalizations show morphological overlap with voice-marked verb forms, and one type shows overlap with the subjunctive gerund forms. Like monomorphemic nouns, many can appear without a possessor or a subject-/object-like element. Conversely, some types impose an interpretive restriction on any genitive item present as the Actor. These features bring up the question of whether some of such instances should be regarded as headless relative clauses. This question will be addressed but will be left open to discussion.

Gerunds also are a kind of nominalization. Indicative gerunds are formed with morphemes similar to those of patient nominalization (§3.6.2) and result nominalization (§3.6.3), namely *-un* and *<n>…-an*. These markers are also homophonous with some of the voice markers: perfective PV *<n>*, non-perfective PV *-un*, and PV/LV *-an*. Subjunctive gerunds are formed via Ce-reduplication, which is also used to create means/manner nominalization. Yet, gerunds diverge from the various types introduced in this section in some key respects. First, the former denote events, while the latter denote entities and concepts. Second, gerunds are fully productive; they can consist of any verb type, and they behave uniformly. Lastly, the subject is obligatory with gerunds.

3.6.1. Actor nominalization

The prefix *me-/mp-*, whose allomorphy is lexically determined, forms actor nominals with the meaning of "one that [verb]-s." These nominalizing prefixes are phonologically identical

with the AV irrealis prefixes, which may encode either habituality or that an event has not yet occurred. The nominalizer *me-/mp-*, too, has the present habitual reading. Due to this homophony, *mp-sapuh* in the following example sentences is ambiguous between a nominalization (3.168a) and an AV-marked verb form (3.168b).

(3.168) mp-sapuh ka Masaw
a. ACTR.NMLZ-practice.medicine NOM Masaw
'Masaw is a doctor.'
b. AV.IRR-practice.medicine NOM Masaw
'Masaw practices medicine.'

At the same time, Actor nominalizations cannot be used to describe a person's occupation or behavior in the past (3.169). It needs to be expressed by means of a verb phrase (3.170).

(3.169) *mp-sapuh seuxal ka Masaw
ACTR.NMLZ-practice.medicine before NOM Masaw
IM: 'Masaw used to be a doctor.'

(3.170) s<m><n>apuh seuxal ka Masaw
practice.medicine<AV><PFV> before NOM Masaw
'Masaw used to practice medicine.'

The possessor[27] of an Actor nominalization bears genitive case (3.171), just as with underived, monomorophemic nouns. Since genitive arguments are prohibited in AV sentences, the mere fact that genitives can co-occur with the *me-/mp*-marked form in (3.171) clearly distinguishes them from (3.172). The latter essentially involves a headless relative clause that means 'one that teaches me.' *Mp-tgsa* acts as an AV-marked verb within the relative clause and, thus, permits an oblique Patient.

(3.171) shiga o, s<n>trgan-an=mu ka mp-tgsa=mu
yesterday TOP meet<PFV>-PV=1SG.GEN NOM ACTR.NMLZ-teach=1SG.GEN
'Yesterday, I ran into my teacher.'

(3.172) shiga o, s<n>trgan-an=mu ka [RC mp-tgsa knan]
yesterday TOP meet<PFV>-PV=1SG.GEN NOM AV.IRR-teach 1SG.OBL
'Yesterday, I ran into the one who teaches me.'

Yet, the combination of *me-/mp-* + verb is not fully productive as a means to derive an Actor nominalization. Consider *me-dudul* in (3.173), which can be analyzed as either a headless relative clause ('one that leads') or as an Actor nominalization ('leader'). The acceptability of the

[27] Note, however, the relationship here is not strictly that of possession. The first person singular possessor is understood to be the one taught by 'the teacher', the possessum.

Patient of *lead* as a genitive possessor, as in *=nami* in (3.173b), depends on the speaker. Therefore, it is perhaps more suitable to analyze these instances as involving relative clauses.

(3.173) a. Masaw ka [$_{RC}$ me-dudul mnan]
Masaw NOM AV.IRR-lead 1PL.EXCL.OBL
'Masaw is the one who leads us.'

b. ?Masaw ka me-dudul=nami
Masaw NOM ACTR.NMLZ-lead=1PL.EXCL.GEN
'Masaw is our leader.'

Nevertheless, it appears that most Actor nominalizations behave like monomorphemic nouns (3.174) in that they can be pluralized with the prefix *d-* (3.175). Note that morphological marking of plurality is optional in Truku, and it seems to be largely restricted to animate nominals. The use of *d-* is archaic but is frequently found in religious texts.

(3.174) d-seediq
PL-person
'people'

(3.175) d-mp-dayaw
PL-ACTR.NMLZ-help
'helpers, servants'

3.6.2. Patient nominalization

Patient nominalizations are marked by *-un*, homophonous with the non-perfective PV suffix, with optional Ce- reduplication. They roughly translate as 'that which is to be [verb]-ed'.

(3.176) u-uq-un
RED-eat-PAT.NMLZ
'food'

(3.177) r-rway-un
RED-admire-PAT.NMLZ
'toy'

The reduplicant is omitted in many cases, though not freely so. It is not clear what conditions regulate this optionality.

(3.178) sinaw o, kingal ?(m-)mah-un
wine TOP one RED-drink-PAT.NMLZ
'Wine is a kind of drink.'

(3.179) ini=ku skuxul (*m-)mah-un s-sibus
NEG=1SG.NOM like.AV.NFIN RED-drink-PAT.NMLZ VBLZ-sugar.cane
'I do not like sweet drinks.'

Patient nominalizations can be used in a sentence without any subject- or object-like items (3.178–180) (cf. the presence of the Actor *=mu* '1SG.GEN' in (3.181). In these cases, forms containing *-un* are treated on a par with monomorphemic nominals.

(3.180) kdal-un=misu [RC p<n>hapuy=mu uq-un]
feed-PV=1SG.GEN:2SG.NOM cook<PFV>=1SG.GEN eat-PAT.NMLZ
'I will let you eat the food that I cooked.'

They can also be quantified with a numeral as in (3.178) and be modified via juxtaposition of a verb phrase (3.179). Note that these characteristics are also compatible with the headless relative analysis.

The genitive element associated with an item marked with *-un* is understood to be the Actor of the base verb. In (3.181), *uq-un* 'food / that which is to be eaten' is intended to be consumed by no one else but the first person singular subject. As discussed in §3.2, Grimshaw (1990) argues that an unambiguous interpretation of the genitive element as subject is evidence for a nominal's argument structure. By this criterion, cases like these may be analyzed either as a Patient nominalization with an argument structure (3.181a) or as a headless relative clause with a PV verb form (3.181b).

(3.181) a. m-n-ekan[28] uq-un=mu ka huling Kuras
AV-PFV-eat eat-PAT.NMLZ=1SG.GEN NOM dog Kuras.GEN
'Kuras' dog ate my food.'

b. m-n-ekan [RC uq-un=mu ka huling Kuras]
AV-PFV-eat eat-PV=1SG.GEN NOM dog Kuras.GEN
'Kuras' dog ate what I was going to eat.'

3.6.3. Result nominalization

Result nominalizations refer to the result or the product of an event (i.e., 'that which has been [verb]-ed') and are formed on activity verbs with *<n>*, which are homophonous with the PV/perfective marker. To this extent, the result nominalizer shares the perfective element with the homophonous voice morpheme.

(3.182) r<n>isuh
draw<RES.NMLZ>
'drawing'

[28] The verb forms for 'eat' in Truku are irregular, containing etymologically unrelated roots *ekan* and *uq* in AV and NAV forms, respectively.

(3.183) n-hapuy
RES.NMLZ-cook
'(prepared) food'

In addition to *<n>*, some result nominals take the *-an* suffix, which is homophonous with the LV/PV marker.

(3.184) k<n>rut-an
cut<RES.NMLZ>-NMLZ
'pieces'

(3.185) q<n>pah-an
work<RES.NMLZ>-NMLZ
'field, garden (where plants are grown)' (lit. 'that which has been worked')

Result nominals show properties similar to patient nominals. They require no co-occurrence of subject- or object-like elements. They can also be modified with quantifiers like *lala* 'many, much'.

(3.186) ekan lala n-hapuy!
eat.AV.NFIN much RES.NMLZ-cook
'Eat a lot of food!'

In addition, the genitive pronoun/noun phrase accompanying a *<n>*-affixed form must be interpreted as the Actor of the base verb. Therefore, *t<n>inun* 'cloth / that which has been woven' in (3.187) refers to a cloth that Hana wove, not a cloth in Hana's possession that was woven by somebody else. To reiterate, with the presence of the genitive Actor, a nominalization analysis (3.187a) and a headless relative analysis (3.187b) are equally plausible.

(3.187) a. malu bi qta-an
good.STAT.FIN INT look-PV
ka m-banah [$_{NP}$ t<n>inun] Hana gaga
NOM STAT.FIN-red weave<RES.NMLZ> Hana.GEN DIST
'That red cloth of Hana's is pretty.'

b. malu bi qta-an
good.STAT.FIN INT look-PV
ka m-banah [$_{RC}$ t<n>inun Hana] gaga
NOM STAT.FIN-red weave<PFV> Hana.GEN DIST
'That red thing that Hana wove is pretty.'

3.6.4. Locative nominalization

Locative nominalizations are formed with *-an*, homophonous with the LV marker, with the resulting nominal meaning 'place for [verb]-ing'.

(3.188) ptas-an
write-LOC.NMLZ
'school'

(3.189) puy-an
cook-LOC.NMLZ
'kitchen'

(3.190) tlng-an
sit-LOC.NMLZ
'chair'

Though the morphology is related to the LV marker *-an*, like patient and result nominals, locative nominals are able to appear without the subject / possessor:

(3.191) s<n>liq-an Kuras ka tlng-an nii
destroy<PFV>-PV Kuras.GEN NOM sit-LOC.NMLZ PROX
'Kuras destroyed this chair.'

Some forms suffixed with *-an* unambiguously refer to entities (3.192, 3.195); corresponding LV readings seem to be lacking. Therefore, meanings intended for (3.193) and (3.196) need to be expressed as AV constructions ((3.194) and (3.197), respectively).

(3.192) tlng-an=mu ka nii (locative nominalization)
sit-LOC.NMLZ=1SG.GEN NOM PROX
'This is my chair.'

(3.193) ?tlng-an=mu ka hini (LV)
sit-LV=1SG.GEN NOM here
IM: 'I sit here.'

(3.194) tluung=ku hini (AV)
sit.AV=1SG.NOM here
'I sit here.'

(3.195) ptas-an=mu ka hini (locative nominalization)
write-LOC.NMLZ=1SG.GEN NOM here
'This place is my school.'

(3.196) ?ptas-an=mu ka sapah Ikung (LV)
write-LV=1SG.GEN NOM house Ikung.GEN
IM: 'I study at Ikung's house.'

(3.197) gisu=ku m-atas sapah Ikung (AV)
PROG.PROX=1SG.NOM AV-write house Ikung.GEN
'I study at Ikung's house.'

Furthermore, although LV verb forms can co-occur with the perfective marker <*n*> (3.198), locative nominals cannot. The interpretation of (3.199) indicates that the combination of <*n*> and *-an* is to be understood as the perfective form of LV. In order to use a locative nominalization

in a past/perfective context, it must be modified by the time adverbial *seuxal* ‘before,’ instead of the perfective affix (3.200).

(3.198) t<n>qi-an=mu ka sapah Ikung
sleep<PFV>-PV=1SG.GEN NOM house Ikung.GEN
‘I have slept at Ikung’s house before.’

(3.199) ?p<n>tas-an=mu ka hini
write<PFV>-LV=1SG.GEN NOM here
‘This is where I have studied before.’ IM:*‘This is my (previous) school.’

(3.200) ptas-an=mu seuxal ka hini
write-LOC.NMLZ=1SG.GEN before NOM here
‘This is my old school.’

3.6.5. Instrument nominalization

Some instrument nominals are derived with the *-an* suffix with optional Ce- reduplication. However, the morphology is not productive and is limited to certain lexical items. Considering their semantic differences, I treat the instances of *-an* discussed so far as homophonous, but independent suffixes: part of the *<n>...-an* complex in result nominalization (§3.6.3), locative nominalizer (§3.6.4), and part of the *Ce-...-an* complex in instrument nominalization (current section).

(3.201) wihi o, kingal (m-)mah-an bgu
spoon TOP one RED-drink-INST.NMLZ soup
‘A spoon is a tool for drinking soup.’

(3.202) sruw o, (t-)tkan-an payay
pestle TOP RED-pound-INST.NMLZ unhulled.rice
‘A pestle is a tool for pounding rice.’

3.6.6. Means/manner nominalization

Ce- reduplication of the verb stem forms a nominal with the semantics of ‘means of / manner of [verb]-ing.’

(3.203) biq-i=ku k-krut=mu damat
give-PV.NFIN=1SG.NOM MAN.NMLZ-cut=1SG.GEN vegetables.OBL
‘Give me something for me to cut vegetables with.’

(3.204) ini=ku skuxul ka t-tgsa=na
NEG=1SG.NOM like.AV.NFIN NOM MAN.NMLZ-teach=3SG.GEN
‘I do not like his way of teaching.’

Even though there is a level of semantic overlap, means nominalization differs from instrumental nominalization in that the latter does not necessarily represent a concrete object such as an instrument. In (3.205), the reason why the speaker cannot achieve his/her goal could

be due to the lack of a cooking tool, time constraints, and so on. The referent of *p-phapuy* is contextually determined.

(3.205) km-phapuy=ku bi siyang nii o,
DES-cook=1SG.NOM INT meat.OBL PROX TOP
ungat ka p-phapuy=mu
NEG.EXIST NOM MAN.NMLZ-cook=1SG.GEN
'I really want to cook this meat, but I cannot.' (lit. 'My means of cooking does not exist.')

In some instances of means/manner nominalization, the possessor / Actor cannot be omitted (3.206). It is not yet clear what conditions its omission. Unlike instrumental nominalization (§3.6.5), this type of nominalization is quite productive. Both its productivity and its subject requirement make means/manner nominalizations difficult to differentiate from other usages of Ce-reduplication, such as subjunctive gerunds and CV future forms (§5.5.2). Alternative glosses are provided in (3.206–207).

(3.206) tama bubu=mu o,
father mother=1SG.GEN TOP
a. ga m-p-rngag quri [NP t-tgsa*(=dha)] knan
PROG.DIST AV-RECP-speak about MAN.NMLZ-teach=3PL.GEN 1SG.OBL
'My parents are talking about my education (lit. their manner of teaching me).'
b. ga m-p-rngag quri [GER t-tgsa*(=dha) knan]
PROG.DIST AV-RECP-speak about SBJV.GER-teach=3PL.GEN 1SG.OBL
'My parents are talking about (their) teaching me.'

(3.207) a. yayu o, [NP k-krut] damat
kitchen.knife TOP MAN.NMLZ-cut vegetable.OBL
'A kitchen knife is a means for cutting vegetables.'
b. yayu o, [GER k-krut damat]
kitchen.knife TOP SBJV.GER-cut vegetable.OBL
'A kitchen knife is for cutting vegetables.'
c. yayu o, [RC k-krut damat]
kitchen.knife TOP CV.FUT-cut vegetable.OBL
'A kitchen knife is what is used to cut vegetables with.'

Yet, Ce-reduplicated forms in instances like (3.207) clearly refer to entities rather than events, defying the gerund analysis.

3.6.7. Abstract nominalization

Many nominals referring to abstract concepts are formed with *kn-* on stative verbs.

(3.208) kn-malu ni kn-naqih
STAT.NMLZ-good and STAT.NMLZ-bad
'good and evil'

(3.209) kn-lala
STAT.NMLZ-many
'quantity'

Some abstract nominals are marked simultaneously by *kn-* and the optional suffix *-an*. The two forms, with or without *-an*, are used interchangeably.

(3.210) baka bi ka kn-sibus(-an) emu nii
enough.STAT.FIN INT NOM STAT.NMLZ-sweet-NMLZ candy PROX
'These candies are just sweet enough.' (lit. 'The sweetness of these candies is just enough.')

While other forms require *-an* (3.211) in addition to *kn-*, there are yet others from which *-an* must be excluded (3.212). This variation seems to be lexically determined as there are no clearly discernible patterns.

(3.211) kn-talux*(-an)
STAT.NMLZ-hot-NMLZ
'temperature, hotness'

(3.212) manu ka kn-malu(*-an)=su?
what NOM STAT.NMLZ-good-NMLZ=2SG.GEN
'What are your strong points?'

Abstract nominals can appear with (3.213) or without (3.214) a possessor.

(3.213) m-shjil ka kn-naqih=su
STAT.FIN-heavy NOM STAT.NMLZ-bad=2SG.GEN
'Your sins are deep (heavy).'

(3.214) ini tuku kn-lala
NEG enough STAT.NMLZ-many
ka t<n>bug-an=mu rudux
NOM feed<PFV>-PV=1SG.GEN chicken
'I do not have enough chickens.' (lit. 'The chickens I feed are not enough in number.')

Table 3.4 summarizes the nominalization affixes introduced in this section, as well as their homophony with verbal affixes.

Table 3.4. Truku nominalization affixes

affix	type of nominalization	voice marking function / other verb forms
me-/mp-	actor nominal	AV irrealis
-un	patient nominal	PV non-perfective
<n>/n- (…an)	result nominal	PV perfective
-an	locative nominal	LV
Ce-…-an	instrumental nominal	---
Ce-	means/manner nominal	CV future subjunctive gerund (voice-neutral)
kn-(…-an)	abstract nominal	---

This section provided a preliminary exploration into various types of nominalization in Truku other than gerunds. I demonstrated that many types of apparent nominalization permit an alternative analysis as verbal or gerundive forms. This is due to their homophonous morphological marking, as well as the interpretation of the genitive-marked element as subject.

Compared to these types of nominalization, gerunds are markedly sentential in nature. Namely, they (i) require the subject (possessor), (ii) are fully productive, and (iii) refer to events rather than entities. Due to partial homophony and ambiguous status of the nominalizations discussed in this section, the clear-cut distinction between derived nominals and gerunds in languages like English (cf. §3.2) cannot be applied to Truku. Nevertheless, I conclude that a comparable contrast does exist in the latter. While both types exhibit co-existing nominal and verbal/sentential properties, gerunds lie on the verbal/sentential end of the spectrum.

3.7. External possession

External possession was briefly introduced in §3.4.3 as a source of optional nominative marking on the gerundive subject. Although its details are not crucial in the characterization of Truku gerunds, external possession itself is a typologically widespread, yet quite diverse, phenomenon that has attracted much theoretical attention. This section provides a more in-depth account for interested audiences.

"External possession" (Payne and Barshi 1999:3, henceforth EP) is a phenomenon in which "a semantic possessor-possessum relation is expressed by coding the possessor…as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum." It is a widespread phenomenon both over vast geographical regions and across language families, practically encompassing the entire globe: North America, Mesoamerica, South America, Australia, Africa, Asia, Southeast Asia, the Pacific, and Europe. However,

external possession manifests itself in quite diverse ways as Payne and Barshi (1999:3) succinctly summarize:

> *The [possessor] may be expressed as subject, direct object, indirect object or dative, or as ergative or absolutive depending on the language type — but not, for example, as an oblique. That is, the [possessor] is expressed like a direct, governed, argument of one of the three universally attested basic predicate types (intransitive, transitive, or ditransitive). In addition to being expressed as a core grammatical relation, in some languages the [possessor] can simultaneously be expressed by a pronoun or pronominal affix internal to the NP containing the [possessum]; but this Genitive-NP-internal coding cannot be the only expression of the [possessor]. Furthermore, the possessor-possessum relationship cannot reside in a possessive lexical predicate such as* have, own *or* be located at *and the lexical verb root does not in any other way have a [possessor] within its core argument frame. Thus, despite being coded as a core argument, the [possessor] is not licensed by the argument frame of the verb root itself — and herein resides the intrinsic fascination of EP constructions.*

Take the following examples from French (2.215) and Japanese (2.216). In French "possessor dative"[29] constructions, the possessor appears as a dative clitic (*lui* '3SG.DET'), while the possessum with a definite determiner stands as a direct object (*la main* 'the hand'). In the paired examples to follow, (a)-sentences are non-EP (i.e., the possessor and the possessum form a constituent as a possessive phrase), and (b)-sentences are external possessor constructions (EPCs). The external possessor in each EPC is bold-faced.

(3.215) *French* (Deal 2013:2)

a.	j'ai	pris	sa	main		
	I-have	taken	his	hand		
b.	je	**lui**	ai	pris	la	main
	I	**3SG.DET**	have	taken	the	hand

'I took his hand.'

In Japanese double-nominative constructions, both the possessor and the possessum take nominative case-marking. The relative ordering of the possessor, the possessum, and the predicate remain identical between (3.216a) and (3.216b). However, an adverb may intervene between the possessor and the possessum in (3.216b), suggesting that the two do not form a constituent.

[29] According to Haspelmath 1999, possessor datives are an areal feature of Europe, found not only in Indo-European languages but also in Basque, Maltese, and Hungarian.

(3.216) *Japanese* (Ura 1996:100)

a. mary-no (*totemo) kami-ga naga-i
   Mary-GEN very hair-NOM long-be
b. **mary**-ga (totemo) kami-ga naga-i
   Mary-NOM very hair-NOM long-be

'Mary's hair is (very) long.'

Just as the phenomenon of external possession is diverse in itself, its analyses also vary. A number of EPCs have been analyzed as "possessor raising." Upon this analysis, the possessor and the possessum underlyingly form a syntactic unit prior to an operation that "raises" the possessor to a higher syntactic position. The purported movement in deriving (3.215b) and (3.216b) are schematized in (3.217) and (3.218), respectively. [30]

(3.217) je **lui** ai pris [$_{DP}$ [$_{DP\_POSSR}$ ___] main]

(3.218) **mary**-ga [$_{DP}$ [$_{DP\_POSSR}$ ___] kami]-ga nagai

An alternative view is to treat EPCs as derivationally unrelated to their non-EP equivalents. Both the possessor and the possessum are generated in their respective surface positions. Existing literature rarely shows agreement on a single analysis; for example, while Croft (1985) and Fox (1981) treat Romance possessor datives as an instance of raising, Tuggy (1980) rejects the raising analysis for Spanish, as does Langacker (1993) for French. Moreover, EPCs within a language can be classified into subtypes, some of which are better accounted for using the raising approach, and others with the non-raising approach. For instance, Ura (1996) claims that, of the double nominative constructions in Japanese, only those in which the possessor and the possessum are in an inalienable relation can be analyzed as resulting from possessor raising (as in 3.218). In those involving alienable relations, the seemingly raised possessor is, in actuality, generated in its original position as a "major subject," which is similar to a topic.[31]

Payne and Barshi (1999) maintain that external possession, in a broader sense, also includes a type of *noun incorporation* which is defined in Baker, Aranovich, and Golluscio 2005 as a "phenomenon in which a nominal that would otherwise bear a grammatical relation to the verb (such as direct object) is expressed not as an independent noun phrase, but rather as a

[30] Also see Oiwa 2016 for a possessor-raising analysis of Truku EPCs.

[31] Detailed discussion of Ura's analysis is ommited here. Roughly speaking, the major topic is taken to be generated in the initial position of a clause with nominative case, whereas the external possessor moves to its final position via a raising operation, and receives its nominative case from the main predicate.

morphological root that is integrated into the inflected verb to form a kind of composite form." Noun incorporation is commonly found in polysynthetic languages, some of which allow possessa to be incorporated into the verb, while coding possessors as verbs' arguments. In the Mapudungun example below, the verb 'run' shows morphological agreement with the first person singular possessor, although the logical subject of the running action is 'my horse'.

(3.219) *Mapudungun*[32] (Baker, Aranovich, and Golluscio 2005:170)
iñche lef-kawell-mün
I run-horse-IND.1SG.SUBJ
'My horse ran away.'

In Truku, a possessive relation between two noun phrases is expressed by juxtaposition, whereby the possessum is immediately followed by the possessor in genitive case. Possessor pronouns attach to possessa as genitive clitics. Although genitive case on noun phrases is morphologically unmarked, pronominal clitics in the first person singular, third person singular, and third person plural have distinct nominative and genitive forms.

(3.220) huling Kuras
dog Kuras.GEN
'Kuras' dog'

(3.221) [laqi snaw]=mu
child male=1SG.GEN
'my son'

In EPCs, the possessor alone surfaces in the pivot position instead of a possessive phrase. On the other hand, the possessum appears after the predicate, immediately before the pivot. The possessum is obligatorily accompanied by a genitive clitic coreferential with the possessor. Compare the EPC in (3.222a) with its non-EP counterpart (3.222b). In (3.222b), the possessive phrase *huling=mu* 'my dog' serves as pivot, as indicated by the nominative marking *ka*. In (3.222a), it is the possessor *yaku* (1SG) that occupies the slot for the pivot. As a result, there is a mismatch between nominative case-marking and voice morphology. In terms of thematic mapping, the AV morphology is expected to correspond to the Agent of the action: 'my dog', rather than 'I'.

[32] According to Baker, Aranovich, and Golluscio (2005:138), Mapudungun is "a language of uncertain genetic affiliation spoken in central Chile and adjoining parts of Argentina."

(3.222) a. wada qduriq huling=mu ka **yaku**
PFV escape-AV dog=1SG.GEN NOM 1SG
b. wada qduriq ka huling=mu
PFV escape.AV NOM dog=1SG.GEN
'My dog escaped.'

In EPCs, voice morphology is indeed required to correspond to the thematic role borne by the possessum. Therefore, the possessor of a Patient can appear in the pivot position if the verb is in PV (3.223), but not if it is in AV (3.224).

(3.223) a. wada=mu paq-un tunux=na ka **Iming**
PFV=1SG.NOM hit-PV head=3SG.GEN NOM Iming
b. wada=mu paq-un ka tunux Iming
PFV=1SG.NOM hit-PV NOM head Iming.GEN
'I hit Iming's head.'

(3.224) a. *wada=ku s<m>ipaq tunux=na ka **Iming**
PFV=1SG.NOM hit<AV> head=3SG.GEN NOM Iming
b. wada=ku s<m>ipaq tunux Iming
PFV=1SG.NOM hit<AV> head.OBL Iming.GEN
'I hit Iming's head.'

When the external possessor is a pronoun, it must be in its independent form with the nominative marker *ka* (3.225a) rather than the nominative clitic (3.225b). Thus, even though the linear placement and the surface case-marking are identical to those of a syntactic pivot, the possessor does not exhibit the syntactic prominency of one.

(3.225) a. wada qduriq huling=mu ka **yaku**
PFV escape.AV dog=1SG.GEN NOM 1SG
b. *wada=**ku** qduriq huling=mu (ka **yaku)**
PFV=1SG.NOM escape.AV dog=1SG.GEN NOM 1SG
'My dog escaped.'

Nevertheless, as previously observed by Tsukida (2009), the nominative-marked external possessor allows for relativization, as expected of a pivot. Thus, 'the child whose toy I broke' in (3.226) is derived from an underlying structure similar to (3.227) in which the 'child' is coded as an external possessor.

(3.226) ga l<m>ingis ka **laqi**
PROG.DIST cry<AV> NOM child
[RC s<n>liq-an=mu rngay-un=na]
destroy<PFV>=1SG.GEN adore-PAT.NMLZ=3SG.GEN
'The child whose toy I broke is crying.'

(3.227) s<n>liq-an=mu rgay-un=na ka **laqi** **gaga**
destroy<PFV>=1SG.GEN adore-PAT.NMLZ=3SG.GEN NOM child DIST
'I broke that child's toy.'

Cross-linguistically, external possession is most frequent with unaccusative verbs, less so with transitive verbs, and least with unergative verbs (Payne and Barshi 1999). Truku PRCs are open to any verb type: transitive (3.223, repeated here as 3.228), unergative (3.229), or unaccusative (3.230).

(3.228) a. wada=mu paq-un tunux=na ka **Iming**
PFV=1SG.NOM hit-PV head=3SG.GEN NOM Iming
b. wada=mu paq-un ka tunux Iming
PFV=1SG.NOM hit-PV NOM head Iming.GEN
'I hit Iming's head.'

(3.229) a. wada qduriq huling=mu ka **yaku**
PFV escape.AV dog=1SG.GEN NOM 1SG
b. wada qduriq ka huling=mu
PFV escape.AV NOM dog=1SG.GEN
'My dog escaped.'

(3.230) a. wada m-huqil huling=mu ka **yaku**
PFV AV-die dog=1SG.GEN NOM 1SG
b. wada m-huqil ka huling=mu
PFV AV-die NOM dog=1SG.GEN
'My dog is dead.'

However, *AV-marked transitive verbs* are an exception to this (2.231). This pattern echoes Payne and Barshi's (1999) observation that EP with "transitive subjects" are extremely rare on the interpretation that a transitive subject coincides with the transitive Actor (Agent).

(3.231) a. *m-n-ekan uq-un=mu huling=na ka **Kuras**
AV-PFV-eat eat-PAT.NMLZ=1SG.GEN dog=3SG.GEN NOM Kuras
b. m-n-ekan uq-un=mu ka huling Kuras
AV-PFV-eat eat-PAT.NMLZ=1SG.GEN NOM dog Kuras.GEN
'Kuras' dog ate my food.'

The rarity of transitive subjects encoded as an external possessor led some, including Baker (1988), to erroneously conclude that they are non-existent. Even so, this is not to say that raising from the transitive subject is impossible. In fact, it has been attested in languages like Korean

and Japanese (Ura 1996), as well as in the Nyikina and Nyulnyul languages of Australia (McGregor 1999).[33]

In some languages, EP occurs only in cases where the relationship between the possessor and the possessum is inalienable (e.g., Baker 1999 for Mohawk, Donohue 1999 for Tukang Besi in which "the possessum must refer to a body part or cosanguineal kin term in order to appear in an EP construction" (pp. 380), Haspelmath 1999 for European languages). Shibatani (1994:462) argues that possessor dative constructions, found in Indo-European languages as well as Modern Hebrew, must involve body parts or else have "a sense of adversity or inconvenience befalling the referents of the dative nominals." When body parts are involved, adversative readings are not required.

(3.232) *Modern Hebrew* (Berman 1982:47)

ima	raxaca	le	**Dan**	et	ha	panim
Mom	washed	to	Dan	DO	the	face

'Mom washed Dan's face.'(lit. 'Mom washed the face to Dan.')

(3.233) *Modern Hebrew* (Berman 1982:48)

ha	tinok	lixlex	**li**	et	ha	xulca
the	baby	dirtied	to-me	DO	the	shirt

'The baby dirtied the shirt on me.'

According to Shibatani, this type of semantic conditioning is not exclusive to EP. It is also observed in adversative passives in Japanese and other Asian languages. He maintains that cross-linguistically, an "extra-thematic argument"–an argument of a verb that is not a direct participant in the predicated event–is more easily incorporated into a sentence if it is highly relevant to the said event.

Surprisingly, neither the inalienability condition nor the affectedness effect seems to hold true for Truku. [34] EP is possible for both (3.223, repeated here as 3.234) and (3.235), former of which involves inalienable possession and the latter alienable possession.

[33] While the former two languages are quite liberal in their application of EPCs to transitive subjects, it is restricted in the latter languages to cases where the subject uses his/her body part instrumentally to perform an action.

[34] Note that Truku does not grammatically distinguish alienable possession from inalienable possession.

(3.234) a. wada=mu paq-un tunux=na ka **Iming**
PFV=1SG.GEN hit-PV head=3SG.GEN NOM Iming
b. wada=mu paq-un ka tunux Iming
PFV=1SG.GEN hit-PV NOM head Iming.GEN
'I hit Iming's head.'

(3.235) a. wada=mu paq-un huling=na ka **Kuras**
PFV=1SG.GEN hit-PV dog=3SG.GEN NOM Kuras
b. wada=mu paq-un ka huling Kuras
PFV=1SG.GEN hit-PV NOM dog Kuras.GEN
'I hit Kuras' dog.'

Yet in (3.235), one may interpret the PR *Kuras* to be adversely affected by the hitting action. However, EP is available to even (3.236), which contains no such connotation. The possessor *Iming* is not in any way affected by 'my hearing her song'.

(3.236) a. b<n>hang-an=mu uyas=na ka **Iming**
hear<PFV>-PV=1SG.GEN song=3SG.GEN NOM Iming
b. b<n>hang-an=mu ka uyas Iming
hear<PFV>-PV=1SG.GEN NOM song Iming.GEN
'I heard Iming's song.'

Below is a summary of the characteristics of Truku EPCs:

a. The possessor is in the nominative position; the possessum occurs within the predicate
b. Genitive clitic on the possessum is obligatory and coreferential with the possessor
c. Voice marking must correspond to the thematic role of the possessum
d. Despite nominative case-marking with *ka*, the possessor cannot appear as nominative *clitic*
e. The possessor can be relativized
f. All verb types are allowed, though EP is unavailable for transitive Actor
g. Possessive relationship can be alienable or inalienable; affectedness constraint is absent

CHAPTER 4

GERUNDS IN FORMOSAN LANGUAGES

This chapter aims to determine if it is possible to trace back origins of Truku gerunds by comparing their morphological patterns with those of gerunds in other Formosan languages. The biggest challenge that this process faces is the fact that gerunds are a severely understudied aspect of Formosan linguistics. Not only are they scarcely mentioned, but the terminology is often ill-defined. Nevertheless, it will be shown that current gerundive morphology is most likely an independent development in Seediq. This is because, although cognates of the morphemes used in Seediq gerund formation exist in many Formosan languages, so far as is known, no other language uses all of them for that purpose. I will first provide a summary of gerunds in Formosan languages as represented in the literature, as well as a short description of gerunds in Tgdaya, another variety of Seediq, based primarily on my own data. While a preliminary investigation into Tgdaya shows patterns similar to Truku, gerunds attested in other Formosan languages bear little to no morphological affinity with those of Seediq. I propose that the indicative gerund forming function of *<n>*, *-an*, and *-un* dates back to as early as Proto-Seediq. It resulted from an extension of the nominalizing function associated with the morphemes along with concurrent loss of thematic orientation.

4.1. Austronesian voice morphology

Before I delve into the survey of purported gerunds in Formosan languages, an introduction to the dual function of Austronesian voice morphology is in order.

4.1.1. Reconstructed Proto-Austronesian verbal paradigm

Today, voice markers in many languages with a Philippine-type voice system are reflexes of the PAn markers found in the independent set reconstructed by Wolff (1973): *<um>, *<in>, *-en ([ən]), *-an, and *Si- (originally reconstructed by Wolff as *i-, but the inclusion of Formosan evidence points to *Si-, cf. Dahl 1973). The infix *<in> indicated perfective aspect (past tense). Only in PV did it mark both voice and aspect as it did not co-occur with the aspectually unmarked (nonpast) PV marker *-en. Table 4.1 shows the PAn verb paradigm proposed by Wolff. A subset of languages in Taiwan and the Philippines, including Seediq, also retain reflexes of the dependent set. Dependent verb forms are mostly used as imperatives and in the position following preverbs including negators. See § 2.1 for the full paradigm of Truku voice morphology.

Table 4.1. Proto-Austronesian voice morphology proposed by Wolff (1973)

		"Active" (AV)	"Direct Passive" (PV)	"Local Passive" (LV)	"Instrumental Passive" (CV)
independent	nonpast	-um-	-en	-an	Si-
	past	-inum-	-in-	-in-an	Si- -in- (?)
future or general action		?[35]	RED-…-en	RED-…-an	?
dependent		Ø	-a	-i	-an (?)
subjunctive ("exhortation")		-a	?	-ay	?

The phenomenon of voice is normally considered a matter of verbal morphology. However, some have suggested that all predicates in Philippine-type languages are nominal, where, all sentences are considered copular and equational (De Wolf 1988, Himmelmann 1991, Naylor 1980). This line of analysis is motivated by the "voice-marked" forms' ability to stand in an argument position without further morphological modification as seen in (4.2) below. In the pair of examples below, the phonetically identical form *b<n>ukuy=dha* in Truku serves as a primary predicate (4.1) and as the sentential pivot (4.2). Under the nominal analysis, (4.1) is treated as an equational sentence. It might accordingly be translated, somewhat awkwardly, as 'This thread is what they tied/their tying'.

(4.1) b<n>ukuy=dha ka gasil nii
tie<PFV.PV>=3PL.GEN NOM thread PROX
'They tied this thread.'

(4.2) m-klay bi ka b<n>ukuy=dha
STAT-tight INT NOM tie<PFV.PV>=3PL.GEN
'They tied the knot tightly.' ('What they tied/their tying is tight.')

4.1.2. Nominalizing functions of "voice markers"

In modern Philippine-type languages, in addition to deriving predicates, *<in>, *-en, *-an, and *Si- (but not *<um>) are widely reflected as lexical nominalizers. Ross (2002:41) provides a summary of the functions of these morphemes in thirteen modern Formosan languages: Saisiyat, Atayal, Seediq, Kavalan, Amis, Tsou, Kanakanavu, Saaroa, Rukai, Thao, Pazeh, Puyuma, and Paiwan, of which all but one, Rukai, possess a Philippine-type voice system. While reflexes of *-en, *-an, *Si-, and *<in> recur as nominalizers in many of the languages, *<um> does so in just two.

[35] Wolff (1973:72) notes that: "[a] question mark indicates an uncertain reconstruction — we know that the category existed, but the shape of the affixes which marked this category is unknown."

(4.3) Number of Formosan languages in which reflexes of PAn voice markers function as voice markers and nominalizers (based on Ross 2002)
   a. <um>: AV marker in 12, nominalizer in 2
   b. -en: PV marker in 8, nominalizer in 7
   c. -an: LV marker in 7, Undergoer Voice marker in 1, nominalizer in 13
   d. Si-: CV marker in 5, nominalizer in 6
   e. <in>: perfective marker in 9, nominalizer in 8

Consequently, ambiguity occurs frequently between verbal (voice inflected) constructions and nominal ones in many of the languages. In Chapter 3, I discussed a number of examples from Truku that can be interpreted either as lexical nominalizations or headless relative clauses. One such example is (4.4). The lexical nominal interpretation ('my food') and the relative clause interpretation ('what I was going to eat') for *uq-un=mu* are represented by the alternative glosses (a) and (b), respectively.

(4.4) m-n-ekan uq-un=mu ka huling Kuras
   a. AV-PFV-eat eat-PAT.NMLZ=1SG.GEN NOM dog Kuras.GEN
   'Kuras' dog ate my food.'
   b. AV-PFV-eat eat-PV =1SG.GEN NOM dog Kuras.GEN
   'Kuras' dog ate what I was going to eat.'

Similarly, L. Huang (2002), Li (2002), and Rau (2002), among others, conclude that noun phrases and verb phrases in some Formosan languages are morphologically indistinguishable from each other.

4.1.3. A note on *-an

Especially striking among Ross's (2002) findings is that reflexes of *-an function as a locative nominalizer in all of the thirteen Formosan languages. According to Blust (2013:394) "[*-an] probably is the most widespread suffix in AN languages. In Philippine-type languages either singly or in combination with other affixes it has a number of distinct functions…" Though its reflexes are most commonly used as an LV marker, they function as a locative nominalizer even in some languages without a Philippine-type voice system. In contrast, "nominalisations formed with a reflex of *-en are far less common" in such languages (Blust 2013:395).

Blust (1996) also reports a number of non-canonical usages of the locative nominalizer *-an. The relationship of the derived nominals to the stem include sources (e.g., PAn *waNiS 'tusk of a boar': *waNiS-an 'wild boar'), instruments (e.g., PMP [36] *tahep 'winnow': *tahep-an

[36] PMP: Proto-Malayo-Polynesian.

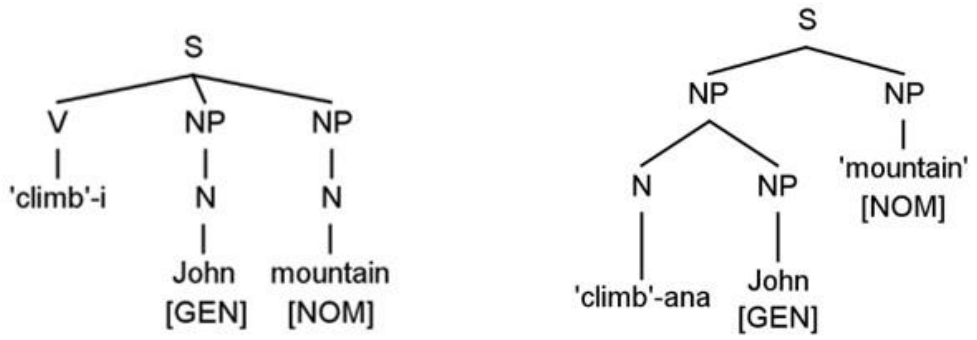
'winnowing basket'), states (e.g., PMP *tian 'abdomen': *tian-an 'pregnant'), and other unclassified relationships (e.g., PAn *bulaw 'golden colored', *bulaw-an 'gold'). Blust's reconstructions suggest that some of these usages date back to either the PMP or the PAn level. As attested in a number of languages in this chapter, reflexes of *-an derive an even wider range of nominals. Such nominals refer to notions that include, but are not limited to, locations, instruments, agents, objects/patients, reasons, temporal expressions, abstract concepts, containers, materials, actions, states, and degrees of states. In at least some of these Formosan varieties, it appears that the suffix has undergone a process of semantic bleaching to become a general nominalizer. This nominalizing function appears to have further extended to gerund formation in some varieties, as well.

4.1.4. Antiquity of "voice markers'" nominalizing function

Starosta, Pawley, and Reid (SPR) 1982 is a seminal work on the voice system of the Proto-Austronesian language. The authors attribute the nominal behavior of voice-marked forms in Philippine-type languages to what they propose as the original function of the morphemes in question: thematic nominalization similar to *-er* and *-ee* in English (e.g., *employ**er*** vs. *employ**ee***). Thematic nominalizers would pick out an argument bearing a particular thematic role from an event, which serves as subject of the sentence (equivalent of *gasil* 'thread' in (4.2)). The resulting nominal (equivalent of *b<n>ukuy=dha* in (4.2), which would be a Patient nominal) would constitute a statement about this argument. Under this analysis, the morphemes' function as voice markers is claimed to be a later development. For SPR, predicates in modern Philippine-type languages can no longer be described as nominal since "while many clauses in languages such as Tagalog, Amis, or Ilokano can be given neat and satisfying analyses as binary NP – NP cleft sentence structures, some can't, due to the fact that a full NP subject occasionally intervenes between the lexical head of the predicate and the other actants of the sentence" (1982:147). Rather than a full-blown voice system comparable to those found in modern Philippine-type languages, SPR maintain that PAn had an "incipient mechanism" (1982:148) of foregrounding various thematic roles. This system was further developed in languages of the Philippines and some languages of Indonesia, Borneo, and Taiwan. Be that as it may, the system "decayed" in the Austronesian-speaking world outside of these areas (SPR 1982:147).

Instead of *<um>, *<in>, *-en, *-an, and *Si-,[37] the "incipient" mechanism in PAn utilized another set of morphemes including *-i and *-a (which correspond to Wolff's "dependent" voice markers) to bring various thematic roles to syntactic prominence. The authors speculate that nominalized forms occurred at such a high frequency in PAn (the proto-language being "strongly noun-oriented…with a high percentage of nominalisation strategies" (1982:149)) that they were eventually re-interpreted as canonical verbal forms. Thematic nominalizations also replaced the original verbal constructions due to surface similarities in word order and case-marking. These similarities are captured in figure 4.1.

Figure 4.1. Structures for PAn sentences with *-i and *-ana (*-an) suffixed forms proposed by Starosta, Pawley, and Reid (1982:157)



'John climbed the mountain.' '[The place where] John climbed is the mountain.'

The original verbal paradigm survived in restricted contexts as evidenced by their continued use as imperative and/or dependent forms in some modern daughter languages. An additional support for this analysis comes from the non-uniformity of the "voice markers," which range from prefixes to infixes and suffixes. SPR note that it is unusual for an inflectional paradigm to display such a high level of morphological diversity, but that this can be accounted for if the morphemes were originally derivational in nature.

Thus, the original voice morphology was replaced by the former thematic nominalizers (now functioning as voice markers). Reflexes of the morphemes in question still function as nominalizers in non-Philippine type languages, leading SPR to place the nominalizer-to-voice marker reanalysis *in the post-PAn times*, although they do not elaborate further on the exact timing of this event. This process of reanalysis is captured by the following PAn sentence.

[37] *-an and *Si- are reconstructed as *-ana and *iSi- in SPR.

(4.5) *Proto-Austronesian* (reconstructed in Ross 2009:302, glosses mine)

	*qaLup-en	ka	babuy	
a.	hunt-PAT.NMLZ	NOM	pig	(before reanalysis)
	'the pigs are something to be hunted'			
b.	hunt-PV	NOM	pig	(after reanalysis)
	'the pigs are hunted' > 's/he hunts the pigs'			

Ross 2002 follows SPR in the spirit of nominalization-to-verb reanalysis. Unlike SPR 1982, Ross (2009) assumes that nominalization-to-verb reanalysis was completed at the PAn level, rather than at the post-PAn level. He does so on the grounds that PAn nominalizations had aspect and mood distinctions, and that it would be unnatural for such distinctions to be absent from verbs within the same language. Thus, "nominalisation and indicative voice were already discrete phenomena in PAn, that is, that derived nominals and indicative verbs were homophonous forms belonging to separate word classes…" (Ross 2002:42).

To date, the most radical nominal analysis of Austronesian syntax is Daniel Kaufman's Nominalist Hypothesis. Kaufman follows SPR's position on the original function of Austronesian voice markers as thematic nominalizers. He departs from SPR 1982 by arguing that Tagalog (and perhaps other "conservative" Austronesian languages, though he does not name any) preserves much of this function (Kaufman 2009). Furthermore, Kaufman claims that not only are Tagalog roots unspecified for lexical category, but the language also lacks the morphological means to verbalize a root.[38] Himmelmann (2008) further observes that Tagalog roots are category-less, since almost all action-denoting roots can appear unaffixed (i.e., with no morphology pertaining to voice or otherwise), and when they do, they obtain an essentially nominal interpretation (state, result, or name of activity). Under Kaufman's view, the reanalysis of nominalizations into verbs postulated by SPR never took place in languages like Tagalog.

In this dissertation, I assume from language-internal evidence that Truku distinguishes the syntactic categories of nouns and verbs. Truku is unlike Tagalog in that action-denoting roots cannot appear unaffixed. Verbs are zero-marked in the nonfinite AV form, whose distribution is restricted to imperative constructions and the position following a preverb. Some verbs are also zero-marked in the finite AV forms, which is a lexically determined pattern. In both cases, I assume an underlying Ø morpheme; they are inflected and not in their bare root forms. Thus,

[38] The issue of parts-of-speech categorization has been raised for other Austronesian languages, most notably Polynesian languages like Samoan and Tongan (Broschart 1997, Mosel and Hovdhaugen 1992).

even if the Nominalist Hypothesis proved to be tenable for some languages, it cannot be applied to all Philippine-type languages.

In the literature on PAn syntax, there is a lack of consensus as to (i) whether thematic nominalization was the only function of *<in>, *-en, *-an, *Si- (and perhaps *<um>) in PAn, and (ii) whether thematic nominalization has been fully reanalyzed as verbal constructions in Philippine-type languages. What is clear, however, is that the morphemes largely have two functions in the daughter languages: voice alternation and lexical thematic nominalization (i.e., the type of nominalization that derives new lexical items).

4.2. A survey of gerunds in Formosan languages

In this section, I will explore constructions described as "gerunds" in existing Formosan literature. However, readers should be aware that their characteristics may not match the definition of gerunds adopted in this dissertation: event-denoting forms exhibiting both prototypically nominal and verbal/sentential characteristics that are derived by a productive process in the syntax. The process is a type of nominalization applied to a verbal stem. In many cases, "gerunds" in the literature are labeled as such with no clear justification. On the other hand, the mere fact that gerunds are not reported in the literature should also not be taken to suggest the lack of the constructions in any given speech variety.

Given that gerunds are not an unusual construction cross-linguistically, their formal characteristics are of particular interest. Namely, Truku gerunds can be divided into two types, which differ in both mood and morphology: indicative and subjunctive. Indicative forms are marked by <*n*>…*-an*, *-an*, or *-un*. When a gerund is marked with <*n*>…*-an*, the event it denotes has a past tense reading. The affix <*n*> indicates perfective aspect elsewhere in the syntax of the language as well. Gerunds marked with *-an* only yield a present tense reading, and those marked with *-un* refer to future events that the speaker deems certain to take place. Recall from Chapter 2 that *-un* marked PV forms are also typically interpreted to refer to events that have yet to occur. These Truku affixes, <*n*>, *-an*, and *-un*, are reflexes of PAn voice morphemes in the finite/independent set: perfective/PV marker *<in>, LV marker *-an, and PV marker *-en, respectively. I argued in Chapter 3 that, despite carrying these morphemes, gerunds are voice-neutral. This was based on the observation that the Actor is the only participant that can serve as a gerundive subject. Therefore, gerunds do not alternate for voice.

Keeping these issues in mind through the sections to follow, I will explore known patterns of gerunds in Formosan varieties other than Truku. For each variety, I will identify how "gerunds" are defined in the relevant sources. Then, the main characteristics of the constructions will be illustrated. I deem further comparisons with Truku gerunds to be unwarranted in cases where the alleged gerunds (i) seem to be particularly limited in productivity or distribution, and/or (ii) do not fit the description of the phenomenon I adhere to in this dissertation. We find that only in Seediq varieties do reflexes of the three PAn voice markers both retain their original function and serve as gerund markers.

4.2.1. Tgdaya Seediq

The Tgdaya variety of Seediq shares much of its verbal morphology in common with Truku, including voice markers. Note that the morphological labels used by Arthur Holmer, who has researched Tgdaya extensively, differ from those I adopted for Truku in this dissertation.

Table 4.2. Tgdaya Seediq voice morphology (adapted from Holmer 2002)

	AV	PV	LV	CV
present	-m- / m-	-un	-an	s-
imperative	ø	-i	-i	-ani
preterit	mn-/-mn-	-n-	-n-an	sn- (?)
immediate future	m-	RED-un	RED-an	s-
distant future	mp-			Ø / p- (?)
irrealis	m- -a	-o	-e	-ane / -ano

Tgdaya has forms affixed with *<n>…-an* that behave in manners comparative to indicative gerunds in Truku created with the same set of markers. Accordingly, I will dub these forms "gerunds." Gerunds marked with *<n>…-an* refer to past events. I have recorded them in both pivot positions (4.6a–c) and adjunct positions (4.6d–e). Note, however, that the nominative marker *ka* is omitted more often in this variety than it is in Truku, rendering it difficult to identify syntactic pivots. The subject of a gerund is the Actor, irrespective of the verb's transitivity (transitive in 4.6b and 4.6d, intransitive in 4.6a, 4.6c, and 4.6e). Thus, just as in Truku, gerunds do not alternate for voice. The subject is consistently marked genitive. As discussed in Chapter 2, genitive case on noun phrases is zero-marked in Truku. In addition, genitive pronominal clitics are only contrastive with their nominative counterparts in the first person singular, the third person singular, and the third person plural. Nominative and genitive are homophonous in other person/number combinations. Whereas Tgdaya uses the same set of pronominal clitics as Truku, genitive case on noun phrases is overtly marked with *na* (e.g., 4.6c), making it easier to observe

case relations. The subject is obligatory, even if its identity is retrievable from the context. For instance, the third person singular clitic *=na* in (4.6e) is coreferential with the pivot of the sentence 'bottle' and cannot be omitted. In addition to the subject, a gerund may include an object bearing oblique case (4.6b).

(4.6) *Tgdaya Seediq*

a. wada teru idas ka n-hqil-an=na di
PFV three month NOM PFV-die-GER=3SG.GEN CS
'He has been dead for three months.' (lit. 'His having been dead is three months.')

b. naqah q<n>lit-an=mu ringo nii
bad.STAT.FIN peel<PFV>-GER=1SG.GEN apple.OBL PROX
'I peeled these apples poorly.' (lit. 'My peeling of these apples is bad.')

c. m-sliko t<n>lang-an na Rabe
STAT.FIN-fast run<PFV>-GER GEN Rabe
'Rabe ran fast.' (lit. 'Rabe's running is fast.')

d. ga l<m>ingis b<n>bey-an=mu ka laqi ga
PROG.DIST cry<AV> hit<PFV>-GER=1SG.GEN NOM child DIST
'The child is crying because I hit him.' (lit. 'The child is crying because of my hitting.')

e. wada ms-peruq p<n>tting-an*(=na) ka rusuq ga
PFV MS-break fall<PFV>-GER=3SG.GEN NOM bottle DIST
'The bottle fell and broke.' (lit. 'The bottle$_i$ broke because of its$_i$ falling.')

The parallelism between (4.6) and the examples of Truku gerunds proves that these constructions are highly akin to each other.

(4.7) *Truku Seediq*

a. wada rima idas ka n-lax-an lumak Kuras da
PFV five month NOM PFV-quit-GER smoke.OBL Kuras.GEN CS
'It has been five months since Kuras quit smoking.' (lit. 'Kuras' having quit smoking has been five months.')

b. naqih hari ka q<n>lit-an=mu supaw nii
bad.STAT.FIN a.little NOM peel<PFV>-GER=1SG.GEN apple.OBL PROX
'I peeled this apple poorly.' (lit. 'My peeling of this apple is a little bad.')

c. m-slikaw ka t<n>lam-an Iming
STAT.FIN-fast NOM run<PFV>-GER Iming.GEN
'Iming ran fast.' (lit. 'Iming's running is fast.')

d. l<m>ingis s<n>liq-an=mu rway-un=na
cry<AV> destroy<PFV>-GER=1SG.GEN adore-PAT.NMLZ=3SG.GEN
ka laqi nii
NOM child PROX
'This child$_i$ is crying because I broke his$_i$ toy.' (lit. 'This child is crying because of my breaking his toy.')

e. l<m>ingis t<n>kur-an*(=na) ka laqi
cry<AV> fall<PFV>-GER=3SG.GEN NOM child
'The child$_i$ cried because he$_i$ fell.' (lit. 'The child cried because of his falling.')

According to Arthur Holmer (pers. comm.), who calls these forms "event nominals," forms affixed with <*n*> alone are also attested. In both Truku and Tgdaya, there is no well-defined set of rules for the optionality of *-an*. Holmer remarks that event-denoting forms with *-un* also exist but are infrequent. I have not been able to record such forms myself. It will be revealed in §5.6.1 that Tgdaya also possesses reduplicated verb forms that resemble Truku subjunctive gerunds in both form and function.

4.2.2. Saisiyat

According to a detailed description of Saisiyat morphology by Elizabeth Zeitoun, Tai-hwa Chu, and Lalo a tahesh kaybaybaw, the language has gerunds that "function as nouns" but "keep verbal properties" (Zeitoun et al. 2015:481). Judging from the English translation provided, many, if not all, instances of gerunds attested in Zeitoun et al. 2015 refer to events (e.g., 'fighting' in 4.9b–c below seems to refer to the act of fighting in general rather than a particular instance thereof), which fits the definition of the category adopted in this dissertation. The authors define the category of gerunds by contrasting it with three other types of nominalization: argument nominalization, action/state nominalization, and clausal nominalization. Following Comrie and Thompson (1985, 2007), the authors take action/state nominalizations to be a process which "[creates] action nouns from action verbs and state nouns from stative verbs or adjectives, meaning the fact, the act, the quality, or occurrence of that verb or adjective" (Comrie and Thompson 1985:350). An argument nominal denotes an argument[39] of the verb it is derived from, which bears thematic roles such as actor, patient, instrumental, locative, and temporal. Nominalized clauses are described as "verbal modifiers," which are essentially relative clauses. The morphosyntactic differences between these four categories of nominalization are listed in table 4.3.

[39] Note that "argument" is used by Comrie and Thompson (1985, 2007) in a loose definition of the term, such that the nominal's syntactic argumenthood to the verb is not presupposed.

Table 4.3. Differences between four types of nominalization in Saisiyat (adapted from Zeitoun et al. 2015)

properties	argument nominalization	action/state nominalization	gerunds	nominalized clauses
contains a nominalizer	yes	yes	no	yes
contains a voice marker	optional	no	yes	yes
occupies an argument position	yes	yes	yes	no
can be modified by another noun	yes	yes	yes	no
may be modified by a possessive pronoun	yes	yes	yes	no
may contain a genitive agent	no	no	no[40]	yes
may co-occur with mood/aspect	yes	no	yes	yes
may co-occur with negation	no	no	no	yes
may be modified by an adverbial (e.g., very, little)	no	no	no	yes
can take an accusative argument	yes	no	yes	yes
may undergo reduplication	yes	yes	yes	yes
may undergo re-verbalization	yes	no	no	no
may undergo re-nominalization	yes	no	no	no

Gerunds occur as subject of certain expressions including *kayzach/'aewhay kita'-en/bazae-'en* 'good/bad to see/hear' (4.8b), as object of verbs like *bazae'* 'to hear', *k<om>ita'* 'see (AV)', *t<om>ortoroe'* 'teach (AV)'" (4.8b–c, 4.9c), and in negative purposive constructions (4.10b–c). Each of the following examples of gerunds is accompanied by its simple declarative counterpart. Note that unlike many Formosan languages, Saisiyat has basic SVO word order.

(4.8) *Saisiyat* (Zeitoun et al. 2015:490, glosses mine)

a. 'ampoa' t<om>o-tatini' ka 'al'alak? (non-gerund)
why <AV>treat.as-old (wo)man ACC young
'Why do you treat the young (person) as an elder?'

b. yako k<om>ita' nisia to-t<m>o-tatini' (gerund)
1SG.NOM <AV>see 3SG.GEN RED-<AV>treat.as-old (wo)man
ka 'al'alak
ACC young
'I saw him treating a young (person) as an elder.'

c. yako k<om>ita' nisia 'a-t<m>o-tatini' (gerund)
1SG.NOM <AV>see 3SG.GEN IRR-<AV>treat.as-old (wo)man
ka 'al'alak
ACC young
'I saw him treating a young (person) as an elder.'

[40] Despite this description, gerunds like (4.8b–c) seem to contain genitive agents.

(4.9) *Saisiyat* (Zeitoun et al. 2015:490, glosses mine)

a. ’oemaw ki ’ataw rengreng ma-sha-shbet (non-gerund)
Oemaw COM Ataw often AV-RED-beat
‘Oemaw and Ataw often fight together.’

b. ma-sha-shbet ’aewhay kita’-en; pa-k-tikot (gerund)
AV-RED-beat bad see-PV CAUS-STAT-afraid
‘Fighting is not good to see; it is frightful.’

c. ka-k-hayza’-an nipon t<om>ortoroe’ (gerund)
REAL-STAT-have-TEMP.NMLZ Japanese <AV>teach
ka ma-sha-shbet ray ka-pa-sha-shbet-an
ACC AV-RED-beat LOC REAL-DYN-RED-beat-LOC.NMLZ
‘In the past, Japanese taught fighting in fighting halls.’

(4.10) *Saisiyat* (Zeitoun et al. 2015:490–491, glosses mine)

a. moyo m-wai’ rini kanman ray taew’an (non-gerund)
2PL.NOM AV-come here 1SG.LOC LOC house
ka-p-tatoroe’-an ka ka-‘al-no-shayshiat
IRR-DYN-learn-LV ACC NMLZ-speak-DAT-Saisiyat
‘You come to my house to learn Saisiyat.’

b. moyo m-wai’ rini kanman ray taew’an (gerund)
2PL.NOM AV-come here 1SG.LOC LOC house
’am=’okik tatoroe’ ka ka-’al-no-shayshiat;
IRR=NEG learn.AV ACC NMLZ-speak-DAT-Saisiyat
’a(m)=m-wai’ ’a-’ialatar nanaw
IRR=AV-come RED-chat only
‘You come to my house not for learning Saisiyat; you have come to chat only.’

c. moyo m-wai’ rini kanman ray taew’an (gerund)
2PL.NOM AV-come here 1SG.LOC LOC house
’okik ’am=tatoroe’ ka ka-’al-no-shayshiat;
NEG IRR=learn.AV ACC NMLZ-speak-DAT-Saisiyat
’a(m)=m-wai’ ’a-’ialatar nanaw
IRR=AV-come RED-chat only
‘You come to my house not for learning Saisiyat; you have come to chat only.’

Besides being able to appear in argument positions (demonstrated in 4.8–9above), gerunds’ nominal properties include the ability to be modified by another nominal. In (4.11), the modifier consists of a relative clause, which is a type of nominalized clause in Zeitoun et al.’s analysis.

(4.11) *Saisiyat* (Zeitoun et al. 2015:486)

[ni baki’ parain kita’-en] ’a-mashashbet pa-k-tikot
GEN grandfather Parain see-PAT.NMLZ IRR-AV.RECP.beat CAUS-STAT-afraid
‘The fight that was seen by Grandfather Parain was frightening.’

Gerunds can be modified with a possessive phrase (4.12). While genitive case marks both NAV Agents and the possessor of a noun phrase in Saisiyat, possessive pronouns and phrases

either mark possessors or serve as possessive predicates. Thus, possessive modification is seen as another manifestation of nominality.

(4.12) *Saisiyat* (Zeitoun et al. 2015:486)

’an=’oemaw	ki	kalih=a	<u>’a-mashashbet</u>	pa-k-tikot
POSS=Oemaw	COM	Kalih=POSS	IRR-AV.RECP.beat	CAUS-STAT-afraid

‘The fight between Oemaw and Kalih was frightening.’

Gerunds’ verbal/sentential quality can be found in (i) containing a voice marker (4.8–11), (ii) being able to co-occur with mood/aspect markers (4.8), and (iii) the presences of an accusative argument (4.8, 4.10). Notice, however, that gerunds can only be marked for AV, not NAV. There is no overt morphological marking for gerunds. Rather, they are homophonous with the canonical AV forms, which are marked with *M-*[41] in the realis and *’a-/’am=* in the irrealis. Based on minimal pairs like (4.8b–c) alone, it is not clear how this mood distinction affects the interpretation of the gerund at this time.

Gerunds in Saisiyat appear to be derived via a fairly productive process and exhibit a combination of nominal characteristics (e.g., argument-like distribution and modification via another nominal or a possessive phrase) and sentential characteristics (e.g., carrying verbal morphology, such as voice and mood/aspect, and taking accusative complements). Forms used as gerunds do not contain gerund markers *per se* but are consistently marked with AV morphology. There are instances of gerunds modified with a genitive or possessive attribute, but unlike with Truku gerunds, they are not obligatory.

4.2.3. Mantauran Rukai

In *A Grammar of Mantauran (Rukai)*, Elizabeth Zeitoun classifies nominalizations in Mantauran into five categories: (i) action/state nominals, (ii) gerunds, (iii) argument nominals (consisting of subjective, objective, instrumental/manner/result, reason, locative, and temporal types), (iv) nominalized clauses (which serve as temporal, conditional, or concessive adjuncts), and (v) abstract nominals (locative, temporal, and reason types).

Zeitoun claims that nominalized forms are “identified as such based on their similar distribution with other nominal arguments…and their sharing of the same morphosyntactic properties” (2007:189). These include (i) occurring in an argument position, (ii) ability to be modified by another nominal (serving as the head of the nominal phrase), (iii) co-occurrence with a genitive (as opposed to nominative) pronoun, (iv) ability to be topicalized, (v)

[41] *M-* is a morpheme that surfaces as alternants <*om*>, *m-*, or *mo-*.

compatibility with plurality marking ([+human]) and a quantifier ([-human]), and (vi) triggering of verbal agreement if [+human].

Like Saisiyat, Mantauran has no morpheme that specifically marks gerunds. Gerunds can be (i) in the root form (i.e., zero-marked, nonfinite in Zeitoun's analysis) (4.13b, 4.14b), (ii) marked as finite (4.13a), or (iii) marked as subjunctive (4.15c). A finite verb form bears a voice marker (active voice is marked by *o-* with allomorphs *om-*, *m-*, and *Ø*) or, if it is stative, the prefix *ma-*. Subjunctive mood is marked by *Ø-* or *m-* on dynamic verbs and by *ma-* on stative verbs. All of the examples presented by Zeitoun are in the active voice.

Gerunds in Mantauran are limited in distribution: they appear in interrogative or pseudo-cleft sentences: "zero-marked [i.e., nonfinite] gerunds usually occur in interrogative sentences introduced by 'who', 'which', 'why'" (4.13b) and pseudo-cleft sentences (4.14b) (Zeitoun 2007:196). Finite-marked gerunds appear in when-interrogative sentences in realis mood (4.15b), whereas those marked subjunctive appear in when-interrogative sentences in irrealis mood (4.15c).

(4.13) *Mantauran Rukai* (Zeitoun 2007:196, glosses mine)

a. o-cengele-ka-'ine? (gerund)
ACT-see-NEG-2SG.GEN-3SG.OBL
'Did you see her?'

b. kani cengele-'-ine (gerund)
why see.NFIN-2SG.GEN-3SG.OBL
'Why did you look at her?'

(4.14) *Mantauran Rukai* (Zeitoun 2007:196, glosses mine)

a. maka-solate-nga ta-ka-eaea solate dhipolo
ACT.finish-write.NFIN-already SUBJ.NMLZ-STAT.NFIN-one.RED book Dhipolo
'Dhipolo has finished writing one book.' (non-gerund)

b. ana solate 'a solate-ni dhipolo (gerund)
that book TOP write.NFIN-3SG.GEN Dhipolo
'That book, it is Dhipolo who wrote it.'

(4.15) *Mantauran Rukai* (Zeitoun 2007:196, glosses mine)

a. ona'i a-kaava'i-ni ana o-dhaace-nga-lrao (non-gerund)
that CLS.NMLZ-come-3SG.GEN that ACT-leave-already-1SG.NOM
'When he came, I had already left.'

b. o-dhaace-'o idhae? (gerund)
ACT-leave-2SG.GEN when.REAL
'When did you leave?'

c. dhoace-'o lo idhae? (gerund)
leave.SBJV-2SG.GEN when.IRR
'When will you leave?'

Based on this description, gerunds in Mantauran appear in specialized constructions only. Though they consist of some sort of nominalization, they are not functionally on a par with Truku gerunds.

4.2.4. Budai Rukai

In the existing Formosan literature, Li-May Sung's article on nominalizations in Budai provides the most comprehensive account of gerunds. According to Sung, Budai gerunds behave much like English Poss-ing gerunds (e.g., 'the doctor's examining the patient'). She clearly states that they constitute a type of *clausal* nominalization, and that they refer to *events* rather than entities. Gerunds are further divided into two types that differ both morphologically and syntactically: *-ane* gerunds and *-Ø* (zero-marked) gerunds. Below is an example of an *-ane* gerund (4.16a), a *-Ø* gerund (4.16b), and a corresponding finite declarative sentence (4.16c).

(4.16) *Budai Rukai* (Sung 2011:525)

a. masamali-aku ku ta-sulrav-ane ki salrabu ki lrailrai
surprise-1SG.NOM OBL NFUT-cure-NMLZ GEN Salrabu OBL Lrailrai
'I am surprised at Salrabu's curing Lrailrai.'

b. masamali-aku ku Ø-sulrav-Ø ki salrabu ki lrailrai
surprise-1SG.NOM OBL NFIN-cure-NMLZ GEN Salrabu OBL Lrailrai
'I am surprised at Salrabu's curing Lrailrai.'

c. wa-sulraw ki lrailrai ka salrabu
FIN-cure OBL Lrailrai NOM Salrabu
'Salrabu cured Lrilrai.'

According to Sung, the gerunds' nominal characteristics include "embeddedness under a case marker /demonstrative and the presence of genitive subject" [42] (2011:538). Meanwhile, their internal structure "resembles in nearly all respects that of a sentence" (2011:541). First, the direct

[42] Budai has three pre-nominal case markers for nouns and noun phrases, *ka*, *ku*, and *ki*, that interact with various semantic factors. The table below summarizes the Budai case-marking system proposed by Sung. According to an alternative analysis by Chen (2008), the prototypical case marked by *ka* and *ku* is nominative and accusative, respectively, while *ki* has a double function of marking oblique and genitive.

Budai Rukai case markers (Sung 2011:528)

Nominative	**Oblique**	**Genitive**
ku (-visible, + distance, ±animate)	ku (-visible, +distance, -human, ±generic)	
ka (+visible, -distance, ±animate)	ka (-visible, +distance, -human)	
	ki (+specific, +human) (+generic, -human)	ki (±animate)

object in (4.17a–b) bears oblique case, just as it does in the finite clause (4.16c). Second, gerunds can contain adverbial phrases like ‘three times’ and ‘yesterday’ (4.18).

(4.17) *Budai Rukai* (Sung 2011:543–4)

a. masamali-aku ku ta-lrumadh-ane ki salrabu ki lrailrai
surprise-1SG.NOM OBL NFUT-hit-NMLZ GEN Salrabu OBL Lrailrai
sangutulrulru kuiya
three.times yesterday
‘I am surprised at Salrabu’s hitting Lrailrai three times yesterday.’

b. masamali-aku ku Ø-lrumay-Ø ki salrabu ki lrailrai
surprise-1SG.NOM OBL ACT-hit-NMLZ GEN Salrabu OBL Lrailrai
sangutulrulru kuiya
three.times yesterday
‘I am surprised at Salrabu’s hitting Lrailrai three times yesterday.’

There are verbal morphemes that *-ane* gerunds can contain, including the passive *ki-* and tense markers (4.18a, 4.19a). Such marking is disallowed in *-Ø* gerunds (4.18b, 4.19b). Note that gerunds (as well as other forms marked with *-ane*, discussed below) choose tense markers from a set different from finite sentences in active or passive voice (table 4.4).

(4.18) *Budai Rukai* (Sung 2011:546–7)

a. masamali ka salrabu ku **ta-ki**-draedrangelr-ane-li
surprise NOM Salrabu OBL NFUT-PASS-cheat-NMLZ-1SG.GEN
‘Salrabu is surprised at my being cheated.’

b. *masamali-aku ku **ki**-lrumay-Ø-ini ki lrailrai
surprise-1SG.NOM OBL PASS-hit-NMLZ-3SG.GEN GEN Lrailrai
‘I am surprised at Lrailrai’s being hit.’

(4.19) *Budai Rukai* (Sung 2011:546–7)

a. lri-kisamula ki-paysu ka salrabu
FUT-work.hard get-money NOM Salrabu
pakai ku **a-ki**-ka-dalam-ane nakuane
for OBL FUT-PASS-STAT.NFIN-like-NMLZ 1SG.OBL
‘Salrabu will/wants to work hard to earn money in order to be liked by me.’

b. *masamali-aku ku **lri-**Ø-lrumay-Ø
surprise-1SG.NOM OBL FUT-ACT-hit-NMLZ-3SG.GEN
ki salrabu ki Lrailrai
GEN Salrabu OBL Lrailrai
IM: ‘I am surprised that Salrabu is going to hit Lrailrai.’

Table 4.4. Tense marking in Budai Rukai (adapted from Chen 2008:31)

	active/passive voice	*-ane* forms
nonfuture tense	*wa-, -a-, Ø, ma- (stative)*	*ta-*
future tense	*lri-/i-*	*a-*

Other verbal morphology, like reciprocal (4.20a), reflexive (4.20b), and causative (4.20c), can also occur within *-ane* gerund forms.

(4.20) *Budai Rukai* (Sung 2011:547)

a. masamali-aku ka ta-**ma**-la-lrumadh-ane
surprise-1SG.NOM OBL NFUT-RECP-RED-hit-NMLZ
ki salrabu si muni
GEN Salrabu and Muni
'I am surprised at Salrabu and Muni's hitting each other.'

b. masamali-aku ka ta-**ngi**-pa-pacadh-ane-ini
surprise-1SG.NOM OBL NFUT-REFL-CAUS-die-NMLZ-3SG.GEN
'I am surprised at his killing himself.'

c. masamali-aku ki salrabu ka ta-**pa**-lrumadh-ane
surprise-1SG.NOM GEN Salrabu OBL NFUT-CAUS-hit-NMLZ
ki muni ki lrailrai
OBL Muni OBL Lrailrai
'I am surprised at Salrabu's making Muni hit Lrailrai.'

Expectedly, Budai gerunds show nominal syntactic properties as well as verbal ones. The nominal negator *kai ka* can appear with either an *-ane* gerund (4.21) or a *-Ø* gerund (4.22), but the verbal negator *kai* cannot.

(4.21) *Budai Rukai* (Sung 2011:548)

a. malisi-aku ka/ ku **kai ka** ta-kela-ane ki salrabu
angry-1SG.NOM OBL OBL NEG KA NFUT-come-NMLZ GEN Salrabu

b. *malisi-aku ka/ ku **kai** ta-kela-ane ki salrabu
angry-1SG.NOM OBL OBL NEG NFUT-come-NMLZ GEN Salrabu

c. *malisi-aku ka/ ku ta-**kai**-kela-ane ki salrabu
angry-1SG.NOM OBL OBL NFUT- NEG-come-NMLZ GEN Salrabu

'I am angry at Salrabu's not coming.'

(4.22) *Budai Rukai* (Sung 2011:548)

a. malisi-aku ka/ ku **kai ka** kela-Ø ki salrabu
angry-1SG.NOM OBL OBL NEG KA come-NMLZ GEN Salrabu

b. *malisi-aku ka/ ku **kai** kela-Ø ki salrabu
angry-1SG.NOM OBL OBL NEG come-NMLZ GEN Salrabu

'I am angry at Salrabu's not coming.'

To summarize, Budai Rukai has two types of gerunds with distinct morphosyntactic behavior. Semantically, gerunds refer to events. They constitute an in-between category which displays both nominal properties and verbal ones. Gerunds' nominality is found in their distribution, genitive case-marking on the subject, and negation with the nominal negator. On the other hand, gerunds also exhibit oblique case-marking on the object and adverbial modification, which are both signs of verbality. Moreover, *-ane* gerunds, but not *-Ø* gerunds, are rich in verbal

morphology, as they may contain passive voice, tense, reflexive, causative, and reciprocal markers.

Of particular interest are the *-ane* gerunds. Under Sung's analysis, this suffix derives two other constructions besides gerunds: lexical nominalization (result nominals), gerunds, and relative clauses. The suffix derives nominals with a number of semantic types often in conjunction with other affixes and/or reduplication. I list just one example of each category described by Sung:

(4.23) *Budai Rukai* (Sung 2011:551–3)
a. place for V-ing: *wa-sulraw* 'cure': *ta-sulra-sulrav-ane* 'hospital'
b. place/container for N: *sabiki* 'betel nuts': *ta-sabik-ane* 'container for betel nuts'
c. a certain day/time for V-ing: *wa-isadha* 'take a rest': *ta-isadhasadha-ane* 'resting day, Sunday'
d. genuinely locally made; related to Budai cultural tradition: *kiping* 'clothing': *ka-kiping-ane* 'traditional clothing'
e. materials for V-ing, something to V with: *wa-sulrav* 'cure': *sa-sulrav-ane* 'medicine'
f. tools for V-ing: *wa-pacase* 'write': *sa-paca-pacas-ane* 'pen, pencil, chalk'
g. part of sth. left from V-ing: *wa-sulraw* 'cure': *sangu-sulrav-ane* 'medical balm left from curing'
h. strong smell of: *wa-sulraw* 'cure': *sangua-sa-sulrav-ane* 'strong smell of medicine'
i. season: *ma-dralrangedrange* 'hot': *kala-dralrangedrang-ane* 'summer'
j. abstract concept: *ma-thariri* 'good, pretty': *tharir-ane* 'goodness, beauty'

The range of unrelated semantics illustrated in (4.23) indicates that, for the purposes of gerunds and result nominal formation, the locative semantics of the PAn *-an is largely lost in Rukai, whereas its nominalizing function persists.

In addition to result nominals and gerunds, *-ane* is also used to derive certain types of relative clauses. Sung regards this behavior of *-ane* as traces of the four-way voice system reconstructed for PAn in a language commonly known to have a two-way active-passive voice system. The author argues that the reflex of LV *-an was "grammaticalized to encompass almost all (undergoer) participants that get relativized" (2011:555).

(4.24) *Budai Rukai* (Sung 2011:534–5)
a. sangu-a-esai ka [$_{RC}$ ni-sinav-ane ka laymay]
smell-FIN-fragrant NOM NI-wash-NMLZ REL clothes
'The clothes that got washed smell good.'

b. ma-lra-aku kane ka
STAT.FIN-hate-1SG.NOM NFIN.eat OBL
[RC lacenge ka ta-aga-ane ki lrailrai]
vegetable REL NFUT-cook-NMLZ GEN Lrailrai
'I hate to eat the vegetables that Lrailrai has cooked.'

c. [RC sa-paysu-ane] ka salrabu
possess-money-NMLZ NOM Salrabu
'Salrabu is one who possesses a lot of money.'

Object voice, a third voice type proposed by Chen (2008), provides further evidence for vestiges of Philippine-type voice alternation in Budai. In asserting the existence of object voice, Chen rejects the active-passive analysis for the language. Compare the active and passive constructions (4.25), with the corresponding object voice construction (4.26).

(4.25) *Budai Rukai* (Chen 2008:24)

a. wa-kane ku babui ka cumai (active voice)
NFUT-eat ACC boar NOM bear
'The bear ate a boar.'

b. ki-a-kane ki cumai ka babui (passive voice)
PASS-NFUT-eat OBL bear NOM boar
'The boar was eaten by a bear.'

(4.26) *Budai Rukai* (Chen 2008:68)

ta-kane-ane ki cumai ka babui (object voice)
NFUT-eat-OV GEN bear NOM boar
'The boar was what the bear ate (at).'

According to Chen, object voice resembles passive voice in bringing semantic focus to the direct object. As with other *-ane* marked forms discussed above, object voice verbs carry the set of tense markers distinct from active or passive voice. Object voice further differs from passive voice in case-marking patterns. The former marks the Agent and the Patient as genitive and oblique, respectively, while the latter marks them as oblique and nominative, respectively. Not only is the Agent of a passive morphologically marked as oblique, but it is also demoted to an oblique syntactic status as evidenced by its omissibility. In contrast, object voice generally requires the Agent. Finally, while passivization can promote either the direct object or the indirect object to the subject status, only the direct object (Patient) can be the subject of an object voice construction[43].

[43] Budai passives and object voice are similar to the two "passives" in languages like Indonesian (Chung 1976) and Balinese, which are also distinguished as "passives" and "objective voice" by Wechsler and Arka (1998). Under Wechsler and Arka's analysis, the latter is a canonical transitive construction, rather than a derived intransitive construction that a passive is.

The validity for Chen's proposal for object voice cannot be determined at this moment. Nonetheless, Budai data shows that a single suffix, *-ane*, (i) has (relics of) a voice marking function (as reflected in its use in object voice and relativization), (ii) serves as a lexical nominalizer, and (iii) derives gerunds: the three functions that the Truku affixes in question, *<n>*, *-un*, and *-an*, display.

4.2.5. Paiwan

Tang 2002 is a description of nominalizations in Paiwan. It is an article constituting part of the *Language and Linguistics* issue on nominalization in Formosan languages. Only a fraction of the article is dedicated to gerunds, whereas the rest of it focuses on a number of result nominals. Result nominals are defined as those "[expressing] the output of a process or an element associated with the process" (Tang 2002:283). Yet, no clear definition of "gerunds" is provided by the author. She simply states that "Paiwan predicates like *pasaLiv* 'wrong' may take a sentential argument," and that gerunds can replace sentential arguments in this environment. Sentential arguments (4.27b) with gerundive ones (4.27c) are distinguished based on case-marking on the Agent. Though the verb takes an AV marker in both (4.27b) and (4.27c), Agent is marked nominative in the former, as it is in an AV sentence (4.27a). On the other hand, the Agent of a gerund in (4.27c) is marked for genitive case. Note that sentential arguments can be in NAV forms (A. Chang 2006), whereas it is not made explicit whether gerunds can inflect for voice or not. In both sentential arguments and gerunds, a temporal expression like *tucu* 'today' may be placed between the verb and the Agent—a trait that Tang takes to be a sign of a sentential structure.

(4.27) *Paiwan* (Tang 2002:329)

a. na-'-em-ayam ti vuvu tua sipaukuz
PFV-examine-AV NOM old.man ACC present
'The old man examined presents.'

b. pasaLiv a '-em-ayam tucu ti vuvu tua sipaukuz
wrong NOM examine-AV today NOM old.man ACC present
'That the old man examines presents is wrong.' (sentential argument)

c. pasaLiv a '-em-ayam tucu ni vuvu tua sipaukuz
wrong NOM examine-AV today **GEN** old.man ACC present
'That the old man examines presents today is wrong.' (lit. 'The old man's examining presents today is wrong.') (gerund)

Like Saisiyat and Mantauran, it appears that Paiwan does not have any overt morphological marking for gerunds. AV-marked forms can be used either as a finite verb or a gerund, and the

two are primarily distinguished by the case-marking on the Agent: nominative in the former, and genitive in the latter. Although "gerunds" seem to have mixed nominal/sentential characteristics, no observations are available beyond case-marking and insertion of temporal expressions.

4.2.6. Amis

In his sketch of Amis, Kazuhiro Imanishi describes the following examples as gerunds. There are two gerund markers: *ka-* and *pi-*, whose choice depends on the verb's class[44]. They alternate with *ma-* and *mi-*, respectively, in finite forms. In (4.28–9), the verbs carry the prefix *pi-* in gerunds (b-sentences), whereas they carry *mi-* in a simple declarative sentence (a-sentences).

(4.28) *Amis* (Imanishi 2009:130; glosses mine)

a. mi-patay cigra tu fafuy
AV-kill 3SG.NOM ACC pig
'He kills/killed a/the pig.'

b. ca ka-gaqay ku pi-patay nigra tu fafuy (gerund)
NEG STAT.NFIN-good NOM NFIN-kill 3SG.GEN ACC pig
'His (way of) killing the pig is/was not good.'

(4.29) *Amis* (Imanishi 2009:82, glosses mine)

a. mi-sa-raka-rakat ku-ra wawa
AV-SA-RED-walk NOM-that child
'That child walks/is walking around.'

b. ca ka-gaqay ku pi-sa-raka-rakat nu-ra wawa (gerund)
NEG STAT.NFIN-gaqay NOM NFIN-SA[45]-RED-walk GEN-that child
'That child's walking around is not good.'

Imanishi recognizes the in-between category nature of the constructions. He states that gerundive clauses do not contain nominative arguments ("topics" in Imanishi's terms). The subject-like attribute of a gerund is marked as genitive instead. In addition, a gerund can be preceded by the nominative case marker *ku* or the common noun classifier *u*. Incongruently, gerunds can also take accusative complements like *tu fafuy* 'pig (ACC)' in (4.28b).

However, Imanishi's analysis of *ka-/pi-* marked forms like those in (4.28–9) as "gerunds" is not shared by Wu (2006), who describes the interpretation of these forms as a manner of performing an action (4.30). In fact, this interpretation is also reflected in the translation of (4.28b) provided by Imanishi (i.e., 'His (*way of*) killing...').

[44] See Wu 2006 for a detailed morphology-based classification of Amis verbs.

[45] According to Imanishi, the prefix *sa-* means "'making, forming, performing' something" (2009:82). However, Wu posits three distinct meanings associated with *sa-*: intensification, 'to form; to create', and 'to appear to; to pretend to' (2006:154).

(4.30) *Amis* (Wu 2006:69)

tata'ak	k-u	pi-palu	aku
big	NOM-CN	NFIN-beat	1SG.GEN

'My way of beating (people) is severe.'

Outside of so-called gerunds, *ka-* and *pi-* generally function as nonfinite markers. *Ka-*/*pi-* forms appear in imperative sentences (4.31) and after *ca'ay* 'not' (4.32b) (Wu 2006:70).

(4.31) *Amis* (Imanishi 2009:131)

pi-patay	tu-ra	fafuy
NFIN-kill	ACC-that	pig

'Kill that pig!'

(4.32) *Amis* (Wu 2006:134)

a.

ma-palu	n-i	ina	kaku
UV-beat	GEN-PPN	mother	1SG.NOM

'Mother beat me.'

b.

ca'ay	ka-palu	n-i	ina	kaku
NEG	NFIN-beat	GEN-PPN	mother	1SG.NOM

'Mother didn't beat me.'

The distribution of *pi-* and *ka-* largely overlaps, albeit differing slightly, between these two environments (table 4.5).

Table 4.5. Amis verb forms in the affirmative declarative sentences, following *ca'ay* 'not', and in the imperative sentences (adapted from Wu 2006:133-138)

	Actor Voice			Undergoer Voice		
				plain		Instrumental Applicative
predicate forms in affirmative declarative sentences	mi-	-um-	ma-	ma-	-en	sa-
predicate forms following *ca'ay*	pi-	ka-…-um-…	ka-	ka-	(ka)…-en	ka-sa-…
predicate forms in imperative sentences	pi-	ka-…-um-…	ka-	-en		sa-…-en

In addition, *ka-/pi-* marked forms serve as a base for further affixation like instrumental applicative *sa-* and locative applicative *-an*[46] (Wu 2006:70).

(4.33) *palu* 'to beat' > *sa-**pi**-palu* 'beat (instrumental applicative)', ***pi**-palu-an* 'beat (locative applicative)'

[46] Note that Wu (2006) adopts an analysis of the Amis voice system with Actor Voice and Undergoer Voice (UV), the latter of which subsumes "plain" UV, Locative applicative, and Instrumental applicative. They correspond to Patient Voice, Locative Voice, and Instrumental Voice in other literature, e.g., Imanishi 2009 and Liu 2011.

Based on the limited data available, it is not clear whether the *ka-/pi-* marked forms described as "gerunds" by Imanishi are simply instances of nonfinite verbal forms, a kind of manner-denoting nominalization, or indeed gerunds (i.e., productive event-denoting nominalization) with a wider array of usages yet to be described. A fundamental challenge comes from nominalizing processes in Amis being scarcely represented in the literature. In any case, the purported gerund-deriving morphology is unrelated between Amis and any of the Formosan languages for which gerunds are reported to exist.

4.2.7. Puyuma

In her grammar of Puyuma, Stacy Teng identifies the following sentences as containing "gerunds" derived by the *-an* suffix. As we will see, however, these constructions are markedly different from Truku gerunds.

(4.34) *Puyuma* (Teng 2007:208)

kaDu	ku=k<in>a-sagar-an		Da	suan
<INTR>there	my.NOM=<PFV>ka-like-NMLZ		INDF.OBL	dog

'My loving of dogs is like that.'

(4.35) *Puyuma* (Teng 2007:208)

tu_i=pasisi-ay	kan	$pilay_i$	Da	Ta-Tuak-an	Da	suan
3GEN=force-TR2	SG.OBL	Pilay	INDF.OBL	RED-kill-NMLZ	INDF.OBL	dog

'Pilay forced him to kill dogs.'

(4.36) *Puyuma* (Teng 2007:372)

ma-uLep=Diya	ki-anger	Datu	ka-sanan-an
INTR-tired=IPFV	<INTR>get-thought	3SG.GEN.NON-SUBJ[47]	KA-stray-NMLZ

'She's still worrying about his possibly getting lost.'

(4.37) *Puyuma* (Teng 2007:209)

wa-aLak	Da	paTungTungan		
go-take	INDF.OBL	drum		
Da	aDi=Diya	b<in>arekep-an	Da	kaLiT
INDF.OBL	NEG=IPFV	<PFV>assemble-NMLZ	INDF.OBL	skin

'Go get a drum that has not been assembled with a skin.'

[47] Puyuma has distinct marking for the possessor within the nominative subject and the possessor in other positions. In addition, unlike in Truku, non-nominative Actor and the possessor are marked differently.

(4.38) *Puyuma* (Teng 2007:210)

kirami=ku=la	Da	ki<a>karun-an
<INTR>start=1SG.NOM=PFV	INDF.OBL	job
Daku	ka-ruwa-an	kikarun
my.OBL	KA[48]-can-NMLZ	work.AV

'I started a job that I am able to do.'

Teng maintains that gerunds "possess both nominal and verbal properties" (2007:208) based on the two tests she employs to determine a construction's syntactic category: choice of negator and co-occurrence with possessive pronouns. Namely, the verbal negator *aDi* is used in gerunds, but not the nominal negator *ameli*. In contrast, gerunds behave like nouns in terms of case-marking. Some gerunds are accompanied by free possessive pronouns (4.36, 4.38), some appear with genitive pronouns (4.34), and yet others with neither (4.35, 4.37). Just as in Saisiyat and Mantauran, genitive case and possessive case are distinguished in Puyuma: "while both nouns and verbs can be preceded by genitive proclitics, only nouns can co-occur with free possessive pronouns" (Teng 2007:66).

These constructions labeled gerunds do not neatly align with those in Truku either functionally or semantically. Gerundive nominals "typically function as relative clauses [modifying a non-Agent head]… or complementation" (Teng 2007:209) as in (4.35–36) and (4.37–9), respectively. While Truku gerunds appear in both argument and adjunct positions, they do not function as modifiers. Furthermore, gerunds in Puyuma do not uniformly refer to events. For instance, (4.34) and (4.36) can be taken to refer to events, but (4.37) and (4.38) clearly do not. It is possible that what Teng describes as gerunds in Puyuma consist of, despite being morphologically identical, two distinct constructions, one of which more closely resembles Truku gerunds. In the absence of further evidence, however, I tentatively conclude that Puyuma gerunds are incomparable with those in Truku.

4.2.8. Kavalan

Chang and Lee 2002 is another article in the special issue of *Language and Linguistics* on Formosan nominalization. The article's primary goal is to distinguish forms with the suffix *-an* from those with the enclitic *=ay*. Unlike past literature that identified both constructions as

[48] According to Zeitoun and Huang's (2000) cross-linguistic overview, *ka-* is used as a stative marker in Formosan languages including Mayrinax Atayal, Paran (Tgdaya) and Truku varieties of Seediq, Mantauran Rukai, and Southern Paiwan. Teng states that its distribution is not so predictable in Puyuma, although it "appears only in imperative, causative, reciprocal, irrealis, and transitive constructions" (2007:179–80).

nominalized ones, the authors treat the former as nominalizations and the latter as nominal modifiers. In their analysis, *-an* affixation results in both what they call gerunds and derived nominals, contrasted as follows: "gerundive nominals are productive, having a regular semantic relationship with their source verbs, and involve clausal structure, derived nominals, on the other hand, are less productive, having a less regular relationship with their source verbs, and instead involve noun phrase structure" (Chang and Lee 2002:351–2). Based on this definition, gerunds in Kavalan have a potential to be something akin to what I have described as gerunds in Truku. Nevertheless, based on the limited data and translation available, I am unable to judge whether the purported gerunds have similar semantic profiles as those in Truku.

The following examples are identified as gerunds.

(4.39) *Kavalan* (Chang and Lee 2002:364)

a. qaqaytisan zaqis-an-na tu paRin
dangerous climb-NMLZ-3SG.GEN ACC tree
'His climbing a tree is dangerous.'

b. nengi taitaen sarekiaw-an-na
good looking dance-NMLZ-3SG.GEN
'Her dancing is beautiful.'

c. nengi ipiran satezay-an-na
good listen sing-NMLZ-3SG.GEN
'Her singing is sweet (lit. good to listen to).'

All the examples provided have genitive attributes, and (4.39a), as well as (4.40) below, contains the accusative-marked[49] complement.

As for gerunds' distribution, Hsieh (2001) observes that it is limited to the subject of a stative verb (4.40) as well as (4.39) above; they cannot serve as a complement (4.41).

(4.40) *Kavalan* (Hsieh 2011:515, glosses mine)

nengi=ti kiala-an na tama-na tu biabas
good=PFV pick-NMLZ GEN father-3SG.GEN ACC guava
'His father's picking guava was good.'

(4.41) *Kavalan* (Hsieh 2011:515 glosses mine)

a. *tayta-an-ku kiala-an na tama-na tu biabas
see-PV-1SG.GEN pick-NMLZ GEN father-3SG.GEN ACC guava
IM: 'I saw his father's picking guavas.'

b. *qapaR-an-ku zaqis-an-na tu paRin
catch-PV-1SG.GEN climb-NMLZ-3SG.GEN ACC tree
IM: 'I caught him climbing a tree.'

[49] The case marker *tu* is glossed as accusative in Chang 1997, 2000a, and Chang and Lee 2002, while it is glossed as oblique in Hsieh 2011. The original glosses for *tu* were unified as ACC here for simplicity.

Moreover, the semantics of so-called gerunds is not elaborated upon in the literature. Their limited distribution is suggestive of the possibility that these forms refer to the manner of an action (e.g., 'the way his father picked guavas' for 4.40) rather than an event[50].

In addition to gerunds, what Chang and Lee (2002) call "action nominals" are formed with suffixation of *-an* (4.42). They refer "to the name of an [activity] designated by the verb" (2002:352).

(4.42) *Kavalan* (Chang and Lee 2002:356–357)

a. nengi sanu-an na sunis a yau
   good educate-NMLZ GEN child LIG that
   'That child's education is good.'

b. qa-sianem=iku tu kerawkaway-an-ku
   think=1SG.NOM ACC work-NMLZ-1SG.GEN
   'I am thinking/worrying about my work.'

On the surface, action nominals appear identical with gerunds. Although the authors claim that an action nominal does not take complements, clearly not all gerunds do either (cf. 4.39b–c). With no other distinguishing features offered, the distinction between the two categories remains unestablished.

Kavalan voice markers are AV *-um-*, *me-*, *Ø*, PV/LV *-an*, and CV (beneficiary/instrument voice) *ti-* (Chang 2000a, Li and Tsuchida 2006). Chang and Lee (2002) assert that the action nominalizer/gerund marker *-an* should clearly be distinguished from PV/LV *-an* as the forms do not predicate anything of either a participant of an event (e.g., Patient) or a location. In addition, to the extent that the genitive attribute of an *-an* marked nominal designates the subject or the possessor in a wider sense, the suffix clearly does not indicate PV or LV. Though not overtly stated in either Chang and Lee 2002 or Hsieh 2011, gerunds also appear to differ from voice-inflected verbal forms in not accommodating aspectual markers.

In sum, Kavalan has nominalized forms affixed with *-an*, some of which are described as gerunds in the literature. Their functional/semantic overlap with Truku gerunds appears limited, but it is worth noting that the three functions of *-an*, voice alternation, thematic nominalization, and (possible) gerundivization, are something that Kavalan shares with both Truku and Budai Rukai (§4.2.4).

[50] I would like to express my gratitude to Yuko Otsuka for pointing this out to me.

4.2.9. Interim summary

Aside from Truku, there are a few Formosan varieties for which gerunds are attested: Tgdaya Seediq, Saisiyat, Mantauran Rukai, Budai Rukai, Paiwan, Puyuma, Amis, and Kavalan. Nonetheless, it is very important to approach the phenomena with caution as authors do not adopt unified criteria for discerning gerunds from other types of nominalization.

To recapitulate, preliminary analysis reveals that gerunds in Tgdaya largely resemble those of Truku in the respects of morphology, distribution, and syntactic behavior. Aside from Tgdaya, only Saisiyat and Budai Rukai were found to have constructions that are sufficiently comparable to Truku gerunds. Amis and Kavalan also offer promising cases, but their semantics seems to point to the manner of an action, rather than events. Main commonalities and differences between the phenomena in these five speech varieties are summarized in table 4.6.

Table 4.6. Morphosyntactic comparisons between gerunds in Seediq and some Formosan varieties

language	morphology	voice alternation	genitive/possessive attribute	accusative/oblique complement
Seediq	*<n>, -an, -en/-un*	no	required	allowed
Saisiyat	AV <*om*>/*m-*/*mo-* (realis) *'a-/'am=* (irrealis)	no	optional?[51]	allowed
Budai Rukai	*-an, Ø*	yes (*-an* only)	required?	allowed
(Amis)	nonfinite *ka-/pi-*	no	required?	allowed
(Kavalan)	nominalizer *-an*	no	required?	allowed

I demonstrated in §3.4.2 that Truku gerunds are voice-neutral, as the choice of the genitive attribute is restricted to the Actor. Interestingly, Budai is the only Formosan variety in which voice alternation within gerunds is known to take place. This is perhaps due to the fact that the very gerund-marking morphemes in every other variety also doubly function as voice markers (at least elsewhere in its syntax, whether gerunds themselves are truly voice-sensitive or not). Because the Budai passive voice marker *ki-* and the nominalizer *-an* (reflex of PAn *-an) do not historically originate in the same paradigm, their co-occurrence does not result in any semantic conflict.

[51] The question marks indicate that co-occurrence between gerunds and genitive/possessive attributes is noted, but its requirement is yet to be determined.

Despite this, morphological marking that appears on purported gerunds is quite diverse across varieties. Gerunds in Saisiyat are described as AV-inflected forms either in realis or irrealis mood, while those in Amis take the form of nonfinite verbs. In Budai and Kavalan, they are formed with a sort of general nominalizer: *-an*.

Based on these findings, I conclude that the gerund forming function of the morphemes found in Truku and Tgdaya can be traced back at least to, but no further than, the Proto-Seediq level. Despite sporadic attestations, gerunds in many Formosan languages remain poorly described. From what is known thus far, no other language has the array of gerund morphemes cognate with those in Seediq. Thus, the function's further antiquity cannot be established. Interestingly, reflexes of PAn *-an derive so-called gerunds, as well as other nominals, in Budai Rukai and Kavalan. Though it is tempting to attribute its gerund-deriving function to the PAn level, evidence is still inconclusive. First, of the two languages, only Budai enjoys an extensive description of gerunds. Second, reflexes of *-an are exceptionally widespread among Austronesian languages, and, unlike reflexes of other voice morphemes, they seem to have undergone a significant amount of semantic change/leveling (cf. §4.1.3). This peculiarity of *-an could have contributed to its development as a gerund marker in Budai (and possibly Kavalan), whereas the function is not restricted to *-an* in Truku.

4.3. Voice markers/thematic nominalizers to gerund markers

Recall that Truku utilizes three of the five PAn "voice markers" *<in>, *-an, and *-en as gerund markers. The degree of overlap cannot be deemed a coincidence. Whether their original function is analyzed as voice-marking or thematic nominalization (cf. §4.1), it is clear that the PAn morphemes alternated to highlight different arguments or event participants like their reflexes do in many of the daughter languages. Thus, I assume that the voice-neutral gerundive function found in Truku was a secondary development. Due to the absence of this form-to-function correspondence outside of Seediq, any attempt to determine the process of such a development would be a futile one. Suffice it to say that the gerund-forming function was a natural extension from the already nominal characteristics of voice-marked (especially NAV) forms / thematic nominalizations. A question remains as to why the Truku CV marker *s-* (< PAn *Si-) does not play a role in gerund formation. It was shown in §3.6 that *<n>*, *-an*, and *-un* all function as thematic nominalizers in the language, yielding result nominalization, locative nominalization, and patient nominalization, respectively. In contrast, *s-* is not used to derive any

lexical nominals. These facts imply that *<n>*, *-an*, and *-un* had a higher potential to be reanalyzed as a general nominalizer.

Pustejovsky (1995), with his generative theory of lexical semantics, argues that the relationship between process/event nominals and result nominals[52] of the same phonetic shape in English is one of logical polysemy, "where the noun seems to have systematically related senses" (1995:31). Logical polysemy "involves lexical senses which are manifestations of the same basic meaning of the word as it occurs in different contexts". For instance, *bank* in (4.43a) denotes an institution, whereas the same word in (4.43b) refers to a building that houses such an institution.

(4.43) a. The bank raised its interest rates yesterday.
b. The store is next to the newly constructed bank. (Pustejovsky 1995:28)

Logical polysemy is distinguished from homonymy whereby "a lexical item accidentally carries two distinct and unrelated meanings" (Pustejovsky 1995:27), as exemplified by the two senses of *bank* in (4.44a) and (4.44b).

(4.44) a. Mary walked along the bank of the river.
b. HarborBank is the richest bank in the city. (Pustejovsky 1995:27)

Some derived nominals in English, especially those pertaining to creation e.g., *creation* and *development*, have three denotations with its own tense/aspect specification (Pustejovsky 1995:94). Whereas nominals do not overtly mark tense and aspect, each of the denotations has a corresponding tensed proposition. The process reading of c*onstruction* in (4.45a) corresponds to the imperfective aspect (4.45b), its result reading (4.46a) to the past tense (4.46b), and its event reading (4.47a) to the present perfect (4.47b).

(4.45) a. John fell from the ladder during the **construction** of the roof frame.
b. John was **constructing** the roof frame, when he fell from the ladder.
(4.46) a. John's **construction** of the roof frame for the house was done yesterday.
b. John **constructed** the roof frame for the house yesterday.
(4.47) a. With the **construction** of the roof frame complete, John can start shingling
b. Now that John **has constructed** the roof frame, he can start shingling.
(Pustejovsky 1995:170–171)

[52] Truku "result nominals" refer to the product of the action/process denoted by a verb and are perfective by nature. For example, *p<n>hapuy* 'food (lit. that which has been cooked)', derived from *phapuy* 'cook', cannot refer to food that *will have been cooked* by sometime in the future. The term "result nominals" is typically used in the literature to refer to derived nominals that refer to entities, which are also a result of an action/process. However, English result nominals such as *examination* and *creation* have no such aspectual specifications as are associated with Truku result nominals.

In Pustejovsky's view, (4.47a) essentially consists of two events, one imperfective (4.45a) and one perfective (4.46a). For a lexical entry with logical polysemy, each of its senses is essential in understanding the meaning of the proposition it appears in, as, for example, the event of construction entails the process of construction taking place and its end product coming into existence. In each of its senses, however, a particular aspect of the event is emphasized.

In a similar fashion, the reanalysis of voice-marked verb forms/thematic nominals as gerunds proposed in this dissertation may be attributed to the ambiguity between a result reading and an (emergent) event reading of the constructions. For instance, the following pairs of sentences contain the phonetically identical form *q<n>lit-an=mu*. The form in (4.48) constitutes a relative clause 'that I peeled' modifying the head noun 'apple'. On the other hand, the preferred interpretation of the form in (4.49) is as part of a gerundive clause, literally 'my peeling of this apple'. At the same time, a relative clause reading is not completely ruled out, albeit being questionable. This is perhaps due to both pragmatic factors[53] and word order. Namely, head-initial relative clauses are preferred over head-final or head-internal ones in Truku (see §2.9). As illustrated in (4.50), a relative clause interpretation is available for both word order types under the appropriate context.

(4.48) naqih hari ka supaw [$_{RC}$ q<n>lit-an=mu] nii
bad.STAT.FIN a.little NOM apple peel<PFV>-PV=1SG.GEN PROX
'This apple that I peeled has gone bad.'

(4.49) naqih hari ka q<n>lit-an=mu supaw nii
bad.STAT.FIN a.little NOM peel<PFV>-GER=1SG.GEN apple.OBL PROX
'I peeled this apple poorly.' (lit. 'My peeling of this apple is a little bad.')
?'This apple that I peeled has gone bad.'

(4.50) a. ini k-hada balay na
NEG STAT.NFIN-ripe INT still
ka supaw [$_{RC}$ q<n>lit-an Iming]
NOM apple peel<PFV>-PV Iming.GEN
b. ini k-hada balay na
NEG STAT.NFIN-ripe INT still
ka [$_{RC}$ q<n>lit-an supaw Iming]
NOM peel<PFV>-PV apple.OBL Iming.GEN
'The apples Iming peeled are not completely ripe yet.'

Possible ambiguity between a relative clause and a gerund is also reflected in the example below. The pivot *en-da-an=mu m-naqih* can be analyzed as a headless relative clause '(that which) I

[53] Sentence (4.49) was elicited in isolation.

have done wrong', as represented by the first gloss (4.51a). Meanwhile, an event reading of the clause along the lines of '(instances of) my doing wrong' would generate a more or less identical semantic interpretation for the sentence. This alternate interpretation is represented by the second gloss (4.51b).

(4.51) lala bi en-da-an=mu m-naqih
a. a.lot INT [RC PFV-do-PV=1SG.GEN AV-do.wrong]
b. a.lot INT PFV-do-GER=1SG.GEN AV-do.wrong
'I have done many bad things.'

Nonetheless, the predicate *lala* 'many' usually requires an object-denoting expression as its subject (4.52a), while *mn-lala* 'many times' takes an event as subject (4.52b).

(4.52) a. *lala bi g<n>sqi-an=na
a.lot int be.late<PFV>-GER=3SG.GEN
b. mn-lala bi g<n>sqi-an=na
x.times-many INT be.late<PFV>-GER=3SG.GEN
'He has been late many times.' (lit. 'His being late has been many times.')

Similarly, the gerund *g<n>rig=su* 'your dancing' in (4.53) is potentially ambiguous with 'what you danced'. A parallel example containing a headless relative clause is presented as (4.54). The use of the demonstrative *gaga* suggests that the preceding clause *t<n>inun Hana* is not a gerund, since gerunds are generally incompatible with demonstratives (cf. §3.4.6.1).

(4.53) naqih qta-an ka g<n>rig=su
bad.STAT.FIN look-PV NOM dance<PFV.GER>=2SG.GEN
'Your dancing (performance) was unsightly.'

(4.54) malu bi qta-an ka [RC t<n>inun Hana] gaga
good.STAT.FIN INT look-PV NOM sew<PFV.PV> Hana.GEN DIST
'That thing that Hana sewed is pretty.'

This sort of ambiguity in Truku occurs only with transitive verbs. Intransitive verbs normally do not inflect for PV or have corresponding result nominals or patient nominals. Therefore, intransitive verbs marked with *<n>*, *-an,* or *-un* only offer a gerundive interpretation (e.g., (4.53)). For this reason, I suspect that the reanalysis from PV/Patient nominalization to gerund took place in relation to transitive verbs. Once the morphemes were reanalyzed as gerund markers and lost their thematic orientation, they were freely applied to intransitive verbs. Hence, the "third" function of Austronesian voice-markers was born. Note that this loss of thematic orientation was selective to the gerundive environment. The morphemes remained productive for the purposes of voice-inflection and thematic nominalization.

Nevertheless, the notion of logical polysemy cannot be directly applied to the ambiguity between thematic nominalizations/voice-inflected forms and gerunds. This is because logical polysemy stems from within lexical entries, i.e., in the semantic representations of lexical items. Gerund formation is a syntactic process, as are productive thematic nominalization and voice alternation. From another perspective, one could argue that the phenomenon offers an attestation of such ambiguity originating in productive processes before it becomes hardwired in the lexicon.

4.4. Conclusion

In this chapter, I presented a possible historical development leading to Truku gerundive morphology whose function was established in Chapter 3. From cross-varietal comparisons, it became apparent that Tgdaya Seediq possesses gerund markers cognate with those of Truku: *<n>*, *-an*, and *-un*. A survey of gerunds attested in the Formosan literature revealed that, not only are constructions parallel to Seediq gerunds few in number, but their morphological marking was also distinct from Seediq.

I proposed that in Seediq, *<n>*, *-an*, and *-un* obtained indicative gerund marking as their third function in addition to their functions as voice markers and thematic nominalizers. This development was a result of syntactic reanalysis, possibly motivated by the ambiguity between the event reading and the result reading of some result nominals or PV-inflected transitive forms. The markers' thematic orientation was eventually lost so that they could be freely applied to intransitive verbs to form event-denoting gerunds. Chapter 5 demonstrates that the subjunctive gerund morphology in Truku emerged through a process similar to its indicative counterpart in spite of their historical sources being unrelated. Together, the two cases present examples of syntactic reanalysis *without* the loss of the original constructions, giving rise to a number of co-existing homophonous forms.

CHAPTER 5

CA- REDUPLICATION IN AUSTRONESIAN AND ITS CONNECTION WITH TRUKU SUBJUNCTIVE GERUNDS

This chapter focuses on reduplication processes in Formosan languages, specifically involving Ca- (initial consonant + fixed vowel /a/) or a similar unit. Upon investigation, both CV- and Ca- patterns of reduplication reveal a wide range of usages across the Austronesian language family, some of which overlap and others seemingly unrelated with one another. Some of the overlapping functions may be attributed to either universal semantic tendencies of reduplication or shared retention from a common ancestor.

I propose that the Ce- reduplication used to create subjunctive gerunds in Truku is a reflex of PAn *Ca- reduplication, whose reconstructed nominalizing function, according to Blust (1998), is instrumental nominalization. In spite of the historical connection, in synchronic terms, the gerund-forming function of reduplication is unique to Seediq (save some inconclusive evidence for Atayal). This suggests that the function was an independent development.

5.1. CV- reduplication and Ca- reduplication in Austronesian

In the following sections, I will summarize contemporary and reconstructed functions of Ca- and CV- reduplication patterns, both of which are widespread in Austronesian languages (Blust 2013). By doing so, I aim to draw inferences about the origins of Ce- (phonetically [Cə-]) reduplication used to form subjunctive gerund forms in Truku. Although other reduplicative patterns are attested in the language family, I will focus on Ca- reduplication and CV- reduplication due to their cross-linguistic frequency, as well as their phonological similarities with Truku Ce- reduplication. In the course of examination, it becomes evident that Ca- reduplication holds higher relevance to the topic of this dissertation.

First, I will provide an overview of the cross-linguistic patterns of Ca- and CV- reduplication within Austronesian languages, especially among those spoken in Taiwan. This will be followed by a summary of works that attempt to reconstruct functions for either of the patterns at the PAn level. The third subsection will investigate the functions of Ce- reduplication in Truku Seediq. Subsequently, they will be compared against similar processes in some of the speech varieties most closely related to Truku: Tgdaya Seediq and Atayal (Mayrinax variety). While Ce- reduplication has multiple functions in Truku, I conclude that the subjunctive gerund formation most likely has its roots in either means/manner nominalization or the future/irrealis-marking

function of reduplication found across Atayalic languages. The reduplicant in these processes can arguably be reconstructed as *Ca- at the Proto-Atayalic level. Although Ross (2009) attributes the future/irrealis marking function of *Ca- to the PAn level, its depth of antiquity remains questionable.

5.1.1. CV- reduplication

Himmelmann (2005:121) states that reduplication is "probably the most pervasive morphological process in western Austronesian languages." Indeed, CV- is a common unit of reduplication both within and beyond Formosan languages. Although the oft-used label "CV-" suggests that the reduplicant is the onset and the nucleus of the first syllable, the template CVC- often co-exists with CV- within the same language. In cases where their functions completely overlap, the two can be seen as allomorphs. Functions of CV- reduplication vary greatly both within and across languages:

(5.1) Functions of CV- reduplication in select Austronesian languages (Blust 2013:424–425)

a. Bunun (Taiwan): durative aspect (e.g., *ma-asik* 'to sweep' : *ma-a-asik* 'keep sweeping'), collectivity (e.g., *ma-bulav* 'ripe, yellow' : *ma-bu-bulav* 'all ripening at once'), and intensity (e.g., *ma-kuis,* 'slender' : *ma-ku-kuis* 'very slender')
b. Tagalog: Future tense (*b<um>ilí* 'to buy' : *bi-bilí* 'will buy') [54]
c. Manam (New Guinea): Plurality on adjectives (e.g., *salaga* 'be long' : *sa-salaga* 'long (pl.)'), continuative, progressive, or perseverative aspects (e.g., *gara-s* 'scrape' : *ga-gara-s* 'is scraping')
d. Pangasinan (Philippines): plural nouns (e.g., *kunáyon* 'relative' : *ka-kunáyon* 'relatives', *nióg* 'coconut' : *ni-nióg* 'coconuts') (Benton 1971:99–100)
e. Thao (Taiwan): (with another, non-reduplicative affix) repetitive actions (e.g., *iŋkmir* 'grasping in the hand': *m-iŋkmir* 'to grasp or knead': *miŋ<m>iŋkmir* 'grasp or knead repeatedly')

Zeitoun and Wu 2006 is a detailed survey of reduplication[55] in thirteen varieties of twelve Formosan languages, primarily based on existing literature. The authors find CV- reduplication to be associated with functions of plural marking on nouns, diminutive, collective/locative,

[54] Note, however, that Reid dismisses this analysis, pointing out that in Tagalog, it is CV:- reduplication (with vowel lengthening) that "marks imperfective or contemplated activity" while CV- reduplication marks "repetitive activity" (2009:242).

[55] In addition to Ca- reduplication and CV- reduplication, the authors also identify C- reduplication (Squliq Atayal), CVC- reduplication, CVV- reduplication (the first syllable CV- with vowel lengthening), full reduplication, CVCV- reduplication (first two syllables), CV.V- reduplication (first syllable and the vowel only of the second syllable), and right-ward reduplication (last syllable of di-, tri-, or quadri-syllabic roots), as well as triplication).

continuous/repetitive aspect, emphatic future, intensification, attenuation, collectivity/quantification, formation of ordinals, and counting of non-human referents. Functions of CV- reduplication are summarized below in (5.2). The numerals in parentheses indicate the number of languages in which each function is found. Note that I list Atayal[56] and Truku in parentheses because the phonetic form of the reduplicant is always [(C)ə-] in these speech varieties. Zeitoun and Wu treat the pattern in Atayal as initial consonant reduplication (C-) and those in Truku as Ce-[57] reduplication. This difference stems from the status of the schwa in each variety. The sound is commonly analyzed as an epenthetic vowel in Atayal (Li 1981, Rau 1992, H.-C. Huang 2006, Lin 2015). The schwa is phonemic in Truku but can also appear as an allophone of all other vowel phonemes as a result of a lenition rule. Of the usages of C- reduplication in Atayal and Ce- reduplication in Truku, Zeitoun and Wu claim that some were derived from Ca- reduplication and others from CV- reduplication. For Squliq Atayal, their judgment was based on comparisons with another variety, C'uli' Atayal. For Truku, the functions were identified according to patterns found in other languages in their study. For instance, six other Formosan speech varieties mark reciprocity with Ca-, while none does so with CV-. Therefore, they conclude that Truku Ce- in this function must have derived from earlier Ca-.

(5.2) Functions of CV- reduplication in thirteen Formosan speech varieties (Zeitoun and Wu 2006)

a. plural (noun): (Atayal), Bunun, Paiwan, Maga Rukai, (Truku), Thao, Tsou (7)
b. collective/locative (place full of X): Maga Rukai, Saisiyat (2)
c. continuous/repetitive aspect: (Atayal), Bunun, Pazeh, Maga Rukai, Saisiyat, Thao, Tsou (7)
d. emphatic future: Truku (1)
e. intensification: (Atayal), Pazeh, Maga Rukai, Mantauran Rukai, Tsou (5)
f. attenuation (+an): Saisiyat (1)
g. collectivity/quantification (stative verbs): Bunun (1)
h. formation of ordinal: Pazeh (1)
i. counting of non-human referents: Puyuma, Siraya (2)

5.1.2. Ca- reduplication

Ca- reduplication also has a wide distribution within Austronesian, and is particularly well-attested among Formosan languages. Blust (2013:425) notes that it "is found in several Formosan and Philippine languages, in Balinese and Chamorro, in various parts of eastern

[56] The variety of Atayal studied in Zeitoun and Wu 2006 is Squliq.
[57] *e* stands for the schwa in the Truku orthography.

Indonesia, and in a limited form in some Oceanic languages such as Motu of southeast New Guinea." The reduplicant is comprised of the onset of the initial syllable and a fixed vowel /a/. In many languages, the reduplicant consists of /a/ only for bases that lack initial consonants. Blust 2013 provides examples from Thao (Taiwan), Ngaju Dayak (Indonesia), and Balinese (Indonesia).

(5.3) Functions of Ca- reduplication in select Austronesian languages (Blust 2013:425–426)

a. Thao: creates instrument nominals from verbs (e.g., *duruk* 'stab' : *da-duruk* 'skewer'), creates numerals for counting human referents (e.g., *tusha* 'two' : *ta-tusha* 'two (of humans)'), marks durative aspect on dynamic verbs (e.g., *mi-lhilhi* 'stand up' : *mi-lha-lhilhi* 'keep standing'), and marks distributivity on stative verbs (e.g., *ma-diplhaq* 'muddy' : *ma-da-diplhaq* 'muddy all over, covered with mud')
b. Ngaju Dayak: "adds an attenuative or qualifying sense", e.g., *hai* 'large': *hahai* 'rather large', *ka-hijaw* 'greenness': *ba-hijaw* 'green': *ha-hijaw* 'greenish'.
c. Balinese: applies to verbs to derive nominals, some of which are instrumental (e.g., *ŋili* 'clean wax from the ear' : *ka-kili* 'earpick') and others non-instrumental (e.g., *ŋəbat* 'spread something out' : *ka-kəbat* 'sago leaf plate'), though the process may be losing its productivity

Like CV- reduplication, Ca- reduplication is associated with a number of unrelated functions. Among the thirteen speech varieties surveyed, Zeitoun and Wu 2006 find the following:

(5.4) Functions of Ca- reduplication in thirteen Formosan speech varieties (Zeitoun and Wu 2006)

a. nominalization: Amis, Paiwan, Pazeh, Puyuma, Saisiyat, Siraya, Thao (7)
b. counting human referents: Amis, Bunun, Sirarya, Thao (4)
c. reciprocal: (Atayal), Paiwan, Puyuma, Maga Rukai, Mantauran Rukai, Saisiyat, (Truku) (7)
d. continuous/repetitive aspect: Amis, Siraya, Thao (3)
e. stative verb formation: Siraya (1)
f. progressive aspect: Puyuma, Siraya (2)
g. future/irrealis: (Atayal), Puyuma, Saisiyat (3) (also Amis, Wu 2006)
h. intensification and quantification (+an): Saisiyat (1)
i. smell of X (*se-Ca-X*): Truku (1), (also *tu-Ca-X* in Thao, Blust 1998)

5.2. Iconicity of reduplication

Notice that many functions of reduplication are shared among several of the speech varieties in Zeitoun and Wu 2006, including plurality/collectivity, continuous/repetitive/progressive aspect, intensification, nominalization, counting of human referents, and reciprocity. Some of these recurring functions such as plurality, continuous/repetitive/progressive aspect, and intensification

are iconic, and are attested widely across languages of the world as functions of reduplication (Moravcsik 1978, Štekauer et al. 2012). Reciprocity can also be seen as a subtype of repetitive actions, i.e., "repeated occurrence of the same event involving the same participants, but with participant roles reversed" (Moravcsik 1978:320). Other functions like nominalization and counting of human referents, though recurring among Formosan, are *non-iconic*. It is worth noting that in many varieties, the nominalizing function of Ca- reduplication is largely limited to instrumental nominals e.g., Saisiyat *t<om>i.ish* 'wipe' : *ta-ti.ish* 'cloth' (Zeitoun and Wu 2006:117), Thao *t<m>iuz* 'to comb' : *ta-tiuz* 'comb (n.)' (Zeitoun and Wu 2006:123), Pazeh *kuxus* 'to shave': *ka-kuxus* 'razor' (Zeitoun and Wu 2006:110). The functions identified in just one or two of the languages tend to be non-iconic e.g., formation of stative verbs and ordinals. At first glance, cases such as "smell of X" appear to be non-iconic. However, similitive marking is found to be a cross-linguistically common function of reduplication in Moravcsik 1978. Moreover, Blust also reports the template *tu-Ca-X* as "smell of X" in Thao e.g., *cumay* 'bear': *tu-ca-cumay* 'smell of a bear' (1998:34). Marking of future tense and irrealis mood is also non-iconic, though Zeitoun and Wu find it in Atayal, Puyuma, Saisiyat, and Truku, a list of languages to which Amis can be added. In her description of Amis verbal morphology, J.-L. Wu (2006:216) argues that Ca- reduplicated forms are in the irrealis mood, thus denoting "events or states that have not taken place or realized, or those that did not take place or realize in the past."

In addition to the matter of iconicity, the phonological proximity between CV- and Ca- also presents a challenge to the historical treatment of these patterns. The next section demonstrates that linguists do explore the possibility that one pattern was derived from the other.

5.3. CV- reduplication and Ca- reduplication in Proto-Austronesian

This section will provide a summary of the existing literature on the historical reconstruction of Ca- reduplication and CV- reduplication at the Proto-Austronesian level. Mainly, there are three positions: (i) Blust, who reconstructs two functions for *Ca- reduplication, neither of which is verbal (instrumental nominalization and formation of human-counting numerals), (ii) Ross, who reconstructs verbal functions of imperfective and future/irrealis for *Ca- reduplication (which is allegedly reflected as CV- in some languages), and (iii) Reid, who does not reconstruct any reduplication, but claims that Ca- was derived from earlier CV-.

5.3.1. Blust

Blust (1999b, 2013) reconstructs two functions of *Ca- reduplication in PAn: (i) the formation of numerals that were used in counting human referents and (ii) the formation of instrumental nouns. Blust (1998:31) schematizes the reduplication pattern as "Ca-CVCVC, where the consonant of the prefix repeats the first consonant of the stem. If the stem begins with a vowel, the reduplicating syllable is simply a-". He also included in his earlier analysis "the formation of certain verb forms" (1998:31) as a function of PAn *Ca-, though he stated that it was not very clear what roles these forms played. He found that they were reflected in modern languages as non-actor focus future tense (Mayrinax Atayal), durative or iterative (Thao), progressive aspect (Tanan Rukai), and future tense (Puyuma). Furthermore, CV-reduplication in Rukai and Tagalog is used to mark contemplated aspect. Blust hypothesizes that CV- in these languages was a result of historical change from Ca-. However, verb formation was dropped in his later reconstruction. In the remainder of this dissertation, I will primarily focus on the verbal usages of Ca- as well as its instrument-nominal function. I will not consider the numeral usage as it has little relevance to the current topic of gerund formation.

According to Blust's reconstruction, PAn *Ca- derived instrument nouns either on its own or with the locative *-an* suffix. From the extensive list of witnesses Blust (1998) provides, I extract below just one example from each of the main witness languages:

Table 5.1. Instrumental nominalization via Ca- reduplication in Austronesian languages (based on Blust 1998)

language	root	verbal citation form	instrumental nominals with Ca- reduplication
Saisiyat (Formosan)	botoeʔ	botoeʔ 'tie'	ba-botoeʔ 'string'
Pazeh (Formosan)	bulas	mu-bulas 'pour water on ground'	ba-bulas 'ladle, dipper'
Thao (Formosan)	capu	c<m>apu 'sweep'	ca-capu 'broom'
Puyuma (Formosan)	kedan	kedan 'whet'	ka-kedan 'whetstone'
Sangir (Philippine)	pegong	ma-megong[58] 'tie round the middle'	pa-pegong 'belt'
Bolaang Mongondow (Philippine)	dagum	mo-dagum 'sew'	do[59]-dagum 'needle'
Tetum (Central-Malayo Polynesian)	baras	baras 'beat with a rod'	ba-baras 'a rod'
Asilulu (Central-Malayo Polynesian)	hala	hala 'shoulder; carry'	ha-hala 'carrying pole'

[58] Stem-initial voiceless obstruents undergo homorganic nasal substitution upon maN- affixation.

[59] The vowel in the reduplicant is *o* because "[at] some period in its history, Bolaang Mongondow … merged *a and *e (schwa) in prepenultimate syllables. At a later period, *e became *o*" (Blust 1998:44).

Blust (1998) uses the reconstructed function of *Ca- to form instrument nominals as an argument against Starosta, Pawley, and Reid 1982, which claims that PAn *Si- was used for that purpose in the proto-language. He reasons that since the two distinct morphological processes could not have had a completely overlapping function, *Si- must have been a verbal affix, rather than a nominalizer. In particular, he argues that *Si- was an Instrumental Voice marker, a function that is retained by its reflexes in many modern AN languages with a Philippine-type voice system.

5.3.2. Ross 2009

Contra Blust, Ross (2009) takes the function of PAn *Ca- reduplication to be primarily a verbal one, while he considers *Si- to have formed instrument nominals. He argues that *Ca- reduplication marked irrealis mood and imperfectivity in the realis mood, the latter of which corresponds to "continuous/repetitive aspect" in Zeitoun and Wu 1996. Ross is uncertain as to whether the irrealis mood marking and the imperfectivity marking were two unrelated functions of *Ca- reduplication, or if the former was "simply a functional extension" (2009:298) of the latter.

There are two problems with Ross's reconstruction. First, imperfective marking is a highly common function of reduplication beyond Austronesian, suggesting that it is an iconic one (Moravcsik 1978, Štekauer et al. 2012). Kiyomi (1993) finds that, of the thirty Malayo-Polynesian languages she investigated, twenty-five use some form of reduplication to express repetition or continuation, five express progressive mood, six express habituality, and one does so to express imperfective aspect. Furthermore, repetition and continuity occur as a reduplicative function at a remarkable frequency across all language groups encompassed in Kiyomi 1993[60]: twelve out of eighteen in Bantu, sixteen out of twenty in Australian, ten out of fourteen in Papuan, and five out of seven in Austroasiatic. In contrast, Kiyomi found future tense marking to be a function of reduplication in just two of the thirty Malayo-Polynesian languages and in none among the other four groups.

Thus, one may not rule out convergence, as opposed to inheritance, as the source of the shared imperfectivity marking function found within Austronesian. Nevertheless, the recurrent correspondence between function and form across these languages is not to be taken lightly. As Reid (2009) suggests, it is of no surprise if some of the languages inherited the function from a

[60] These groups are called "language families" in Kiyomi 1993. However, note that both Australian and Papuan languages consist of language families and isolates that are not proven to be related to one another.

common ancestor. Also problematic is the fact that Ross considers *Ca- to be "reflected as *Ca-* in Puyuma, Kanakanavu, Saaroa, Thao, Amis and Siraya," but to have been "replaced by CV- in Saisiyat, [Pazeh], Bunun, Paiwan," and Proto-Malayo Polynesian (2009:299). In other words, his reconstruction was based on a conflation of synchronic Ca- and CV- patterns. By reconstructing an iconic function of reduplication from two different forms without the support of regular sound correspondences, he further overlooks the possibility of resemblance by convergence.

Moreover, support for his reconstruction of irrealis marking as a function of PAn *Ca- is weakened by the fact that the languages that purportedly reflect it display quite different paradigmatic patterns. These languages are Amis, Atayal, Puyuma, Saisiyat, and Seediq. Though Amis is not used as evidence for PAn reconstruction in Ross (2009), Wu 2006 reports that it also utilizes Ca- reduplication for irrealis marking. Note that these five languages are genetically diverse, coming from four different primary branches of Austronesian per Blust's (1999a) subgrouping: Amis from the East Formosan subgroup, Saisiyat from the Northwestern Formosan subgroup, and Atayal and Seediq from the Atayalic subgroup, whereas Puyuma forms a branch on its own.

In Amis, while Ca- reduplication occurs in AV- and UV- (≈PV) inflected forms, it does not occur in LV and CV (applicative forms, in Wu's 2006 analysis) forms.

(5.5) *Amis* (J.-L. Wu 2006:216)

a. pa-palu-en n-i sera Ø-ci kuyu
IRR-beat-UV GEN-PPN Sera NOM-PPN Kuyu
'Sera will beat Kuyu.'

b. ma-mi-nanum kaku, mi-tapadang kisu
IRR-AV-water 1SG.NOM AV-call 2SG.NOM
'When I was about to drink water, you called me. (So I didn't drink.)'

c. ma-ma-palu n-i aki Ø-ci panay, piyoc mi-laliw
IRR-UV-beat GEN-PPN Aki NOM-PPN Panay soon AV-run.away
'When Panay was about to be beaten by Aki, she ran away quickly (and did not get beaten.)'

In Atayalic languages, the reduplicant (Ca- in Mayrinax Atayal, Ce- in Truku Seediq, and Cu- in Tgdaya Seediq) co-occurs with reflexes of PAn *-en (PV) and *-an (LV), but not with reflexes of the AV *<um> or the CV *Si-. AV emphatic future verbs in Seediq are in the form of RED-IRR-STEM (e.g., Truku *me-me-iyah* < *me-iyah* 'come (AV.IRR)'). In Mayrinax Atayal, AV forms are *not* reduplicated (L. Huang 2002). In both Seediq and Atayal, CV future/irrealis verbs are in simple Ca- reduplicated forms and carry no overt voice morphology (e.g., Truku *hadut* >

h-hadut ‘send (CV.FUT)). On the other hand, irrealis forms in Puyuma (Ulivelivek, Katipul, and Tamalakaw dialects) are listed in Ross (2009) as *Ca*-STEM (AV), *Ca*-STEM-*en* (PV), *Ca*-STEM-*an* (LV), and *i*-*Ca*-STEM (CV). The Nanwang dialect also has *Ca*-STEM as the AV form, whereas PV and LV take the form of *Ca*-STEM-*i*, and CV takes the form *Ca*-STEM-*an* (Teng 2007). Saisiyat irrealis verbs are quite disparate in form: AV *’am*[61]=STEM<*om*>, PV *ka*-STEM-*en*, LV *ka*-STEM-*an*, and CV *Ca*-STEM (Zeitoun et al. 2015). If irrealis marking of Ca- in Saisiyat *was* a direct continuation from PAn, as Ross claims, it would suggest that the language replaced it with another marker in all voice types other than CV. It would also mean that, while the imperfective usage of Ca- reduplicated verb forms is preserved in many languages, the future/irrealis usage has been lost in most.[62] Both of these hypothetical situations would beg for an explanation.

Table 5.2. Compatibility of C(a)- reduplication for irreais/future marking with different voice types in five Formosan languages

language		AV	PV	LV	CV
Amis	compatible with Ca- (irrealis/future)?	yes	yes	no	no
	voice morphology	marked	marked		
Saisiyat	compatible with Ca- (irrealis/future)?	no	no	no	yes
	voice morphology				zero-marked
Puyuma	compatible with Ca- (irrealis/future)?	yes	yes	yes	yes
	voice morphology	zero-marked	marked	marked	marked
Atayal	compatible with Ca- (irrealis/future)?	no	yes	yes	yes
	voice morphology		marked	marked	zero-marked
Seediq	compatible with Ca- (irrealis/future)?	yes	yes	yes	yes
	voice morphology	marked	marked	marked	zero-marked

[61] Yeh (2006) analyzes *’am=* in this construction as resulting from grammaticalization of the verb *’am* ‘want’.

[62] Continuous/repetitive aspect marking is carried out via CVC- reduplication in Saisiyat (Zeitoun et al.2015). Ca- reduplication marks progressive aspect in addition to irrealis mood in Puyuma, but the two are in complementary distribution. When Ca- marks progressive mood, verb forms carry voice markers (AV <*um*>, PV *-aw*, LV *-ay*, CV *-anay*) different from the set used in the irrealis mood forms. Therefore, the AV progressive form for *kasu* ‘to bring’ is *k*<*em*>*a-kasu*, while its irrealis counterpart is *ka-kasu* (Teng 2007:55).

5.3.3. Reid 2009

Unlike Blust and Ross, Reid (2009) contends that neither CV- reduplication nor Ca- reduplication should be reconstructed for PAn, although some of their functions in the modern Austronesian languages are *probably* inherited from the proto-language. This is because "it is impossible to show that highly iconic reduplications occurring in related languages … are the result of inheritance rather than convergent development" (2009:240). He further argues that even non-iconic reduplicative patterns should not be attributed to cognacy, because "common paths of semantic change" (2009:240) from more iconic meanings can lead to convergence. Having established this position, Reid claims that Ca- reduplication was derived from earlier CV- reduplication rather than vice versa. As supporting evidence, Reid cites Niepokuj's (1997) cross-linguistic study which found that, in all languages with both fixed-vowel reduplication and copy-vowel reduplication, the latter was older than the former. Her evidence consisted of attestation in earlier written texts and productivity; namely that copy-vowel reduplication was less productive than Ca- reduplication.

Validating the reconstructibility of either Ca- reduplication or CV- reduplication and their functions at the PAn level is beyond the scope of this dissertation. Under close scrutiny, Truku Ce- reduplication reveals not only functions shared by other Formosan languages but also some unique ones. Notably, its role in subjunctive gerund-formation seems to be a language-internal development. At the same time, the function can be seen as an extension of either instrumental nominalization or future/irrealis marking. As we saw in the current section, the former is productive in a number of Austronesian languages, and the latter is found in a few Formosan languages.

5.4. Reduplication in Truku

This section provides an illustration of reduplication in Truku. Tsukida's (2009) grammar lists a number of processes, which can be classified on a formal basis as follows: (i) C_1e-/C_1eC_2e-, (ii) Ce-, and (iii) Ce- + additional affix.[63] In the C_1e-/C_1eC_2e- type, the first syllable's onset and nucleus are repeated if the base is trisyllabic or longer. If the base is disyllabic, the onset and nucleus of the second syllable[64] are repeated in addition to the first syllable (2009:181).

[63] As stated earlier, *e* in Truku orthography normally stands for the schwa. The schwa is not spelled out where its existence can be predicted based on the context.

[64] Tsukida (2009) states that the unit of reduplication is the syllable, but that coda consonants are ignored. Since native Truku vocabulary seems to lack root-medial coda consonants and most roosts are at

In Ce- reduplication, the reduplicant invariably consists of the first syllable (e.g., *geuguy* /ɣə'uɣuy/ ‘steal’ > *g-geuguy* [ɣəɣə'uɣuy]). If the stem begins in a vowel, the reduplicant consists of the vowel only (e.g., *iyah* /'iyaχ/ ‘come’ > *e-iyah* [ə'iyaχ]). Some functions of Ce-reduplication require a co-occurring affix. In all reduplication patterns, repeated vowels are realized as the schwa. This is due to the phonological constraint in Truku that reduces most pre-penultimate vowels to the schwa. For instance, /i/ in the penult of *l**i**ngis* ‘to cry’ is realized as [i] in its AV form *l<um>ingis* /ɮum**i**ŋis/ [ɮu'm**i**ŋis]. On the other hand, the same vowel surfaces as [ə] in the verb’s LV form, since it moves to the antepenult as a result of suffixation: *lngis-an* /ɮ**i**ŋisan/ [ɮ**ə**'ŋisan]. The same constraint holds for the Toda and Tgdaya varieties of Seediq, except the reduced vowel is [u] rather than [ə] in the latter.

5.4.1. C_1e-/C_1eC_2e- reduplication

According to Tsukida 2009, C_1e-/C_1eC_2e- reduplication has two functions. It is used optionally to mark plurality on nouns.

(5.6) Plurality marking on nouns (Tsukida 2009:181)
a. laqi ‘child/children’ : lq-laqi [ɮəqə'ɮaqəi̯] ‘children’
b. rudan ‘elder/elders’ : rd-rudan [l̥ədə'l̥odan] ‘elders’
c. qhuni ‘tree/trees’ : q-qhuni [qəqə'āoni] ‘trees’

It also optionally marks plurality on stative verbs.

(5.7) Plurality marking on verbs (Tsukida 2009:266–267)
a. bilaq ka laqi=mu
STAT.FIN.small NOM child=1SG.GEN
‘My child is small.’ / ‘My children are small.’
b. bl-bilaq ka laqi=mu
PL-STAT.FIN.small NOM child=1SG.GEN
‘My children are small.’

5.4.2. Ce- reduplication

5.4.2.1. Future tense marking

Tsukida analyzes the function of stand-alone Ce- reduplication as marking of emphatic future tense. It is applied to the verb stem, which minimally consists of a verb root and a voice marker, but can include other morphemes like the causative prefix *p*-. Tsukida lists the following three

---

least disyllabic, codas play no role in the reduplication of the initial syllable, which is always open. Infixation of the perfective/PV marked <*n*> can introduce the word-medial coda, which consists of a nasal that is homorganic with the following plosive sound e.g., s<n>but [sə**m**.but] ‘beat (PV.PFV)’ > *sbut* /səbut/, q<n>driq-an [qə**n.**də.liɋ.qan] ‘escaping (GER)’ > qduriq /qəduliq/. However, infixes do not phonologically interact with reduplication since the two do not co-occur.

reduplicated forms in her verbal paradigm: (*me-)me*[65]-STEM for Actor Voice, (Ce-)STEM-*un* for Goal Voice (≈Patient Voice), and (Ce-)STEM for Circumstantial Voice. However, it appears that reduplication is infrequent in these environments. Tsukida provides one sentence containing the reduplicated AV irrealis form (5.8). Although PV forms like *p-p-hqil-un* > *p-huqil* 'kill (CAUS-die)' and *q-qta-un* > *qita* 'see, look' are mentioned, she does not provide them in sentences. Moreover, both in her experience and mine, language consultants often remark that reduplication in this context is optional, perhaps due to its emphatic nature.

(5.8) **m**-me-iyah=ku o,
RED-AV.IRR-come=1SG.NOM TOP
jima=ku m-n-ekan n-hapuy sapah=mu
already=1SG.NOM AV-PFV-eat RES.NMLZ-cook house.OBL-1SG.GEN
'When I was about to come, I had already eaten [a] meal at home.'
(Tsukida 2009:810)

I have found a few examples in other sources, including my own field notes, the Bible, and the online dictionary of Truku compiled by the Center of Indigenous Development, National Dong Hua University.

(5.9) lngug=ku **m**-mowsa, asi lu m-narux tunux
intend=1SG.NOM RED-go.IRR EPIS suddenly STAT.FIN-sick head
'I intended to go, (but) I suddenly got a headache.' (field note)

(5.10) **m**-me-dumul=ku m-tahu do
RED-AV.IRR-make.fire=1SG.NOM AV-make.fire CS
asi lu maq ka quyux,
EPIS suddenly dump.mud NOM rain
ki n-alax=mu[66]
so PV.PFV-give.up=1SG.GEN
'When I was about to make fire, rain suddenly started to pour, so I gave up.'

(5.11) (**e**-)uq-un=mu ka bunga nii
RED-eat-PV=1SG.GEN NOM yam PROX
'I will eat this yam.' (field note)

(5.12) r<m>uba sunan ka seediq o, **r**-rba-un ni,
curse<AV> 2SG.OBL NOM person TOP RED-curse-PV and
s-m-bliq sunan ka seediq o, **s**-s-bliq-un
CAUS-AV-blessed 2SG.OBL NOM person TOP RED-CAUS-blessed-PV
'Those who curse you will be cursed, and those who bless you will be blessed.'
(*Soyang Patas*, Genesis 27.29)

[65] *me-* =AV.IRR.

[66] Retrieved on 12/08/16 from 原住民族語言線上詞典 (http://e-dictionary.apc.gov.tw/trv).

More commonly, *Ce-…-un* is found in what appear to be headless relative clauses. I argued in (§3.6.2) that, alternatively, some of these patterns can be analyzed as nominalizations rather than headless relative clauses. From the semantics, the function of the prefix may be more accurately described as marking of irrealis mood rather than future tense.

(5.13) wada=mu kdu-un m-atas
PFV=1SG.GEN finish-PV AV-write
ka [saw **p**-ptas-un=mu] da
NOM SIM RED-write-PV=1SG.GEN CS
'I have finished writing my thesis (thing for me to write).'

(5.14) mp-tgsa lala bi
AV.IRR-teach a.lot INT
[**k**-kla-un=su / **s**-slhay-un=su] ka hiya
RED-know-PV=2SG.GEN RED-learn-PV=2SG.GEN NOM 3SG
'He/she will teach you a lot of things for you to know/learn.'

Tsukida argues that CV verbs in future tense are formed with Ce- reduplication; that is, with no overt voice affix. The pivot of a sentence with a CV-inflected verb can be an instrument (5.15), a beneficiary (5.16), a theme of a transfer verb (5.17), or a reason (§2.1.1). Reduplication is optional, so the verb can be phonetically identical with the verb stem.

(5.15) **b**-bsqar=su manu ka kana puniq gaga?[67]
RED-shoot=2SG.GEN what.OBL NOM all gun DIST
'What are you going to shoot with all those guns?'

(5.16) **p**-p-huqil=mu rudux ka dangi=mu
RED-CAUS-kill.CV=1SG.GEN chicken.OBL NOM fiancé=1SG.GEN
'I will kill a chicken for my fiancé.' (field note)

(5.17) (**r**-)rngag=na laqi-an ka kari nii
RED-tell.CV=3SG.GEN child-OBL NOM word this
'He/she will tell a/the child about this story.' (Tsukida 2009:234)

5.4.2.2. Subjunctive gerund formation

In Chapter 3, I illustrated the characteristics of subjunctive gerunds, which are also constructed via Ce- reduplication. Gerunds constitute an in-between category that displays both nominal and sentential characteristics. In Truku, they can appear in the sentential pivot position, the oblique argument position, the predicate position, adjunct positions, and as the object of prepositions. Subjunctive gerunds, in particular, denote hypothetical events or states.

Tsukida analyzes these patterns as the "nonfinite usage" of CV future forms (2009:745). On the contrary, they can hardly be called CV, being *unmarked* for voice (voice-neutral), as I argued

[67] Retrieved on 11/28/16 from 原住民族語言線上詞典 (http://e-dictionary.apc.gov.tw/trv).

earlier. Therefore, despite being formally identical, subjunctive gerunds and CV future forms can have seemingly contradictory semantics. Compare (5.18), a pseudo-cleft sentence with a CV future verb form, with (5.19), in which a gerund serves as predicate.

(5.18) CV future (Tsukida 2009:235)

hii=mu	ka	**t-talang**=mu	paah	saman
body=1SG.GEN	NOM	CV.FUT-run=1SG.GEN	from	next.morning

Lit. 'The reason I will run from next morning is my body.'
'I will begin jogging from tomorrow morning in order to maintain my health.'

(5.19) Subjunctive gerund

k-k-brax	hii=mu	ka	t<m>alang=ku
SBJV.GER-STAT.NFIN-strong	body.GEN=1SG.GEN	NOM	run<AV>=1SG.NOM

'I run for my health.' (lit. '(The reason why) I run is for my body's being strong.')

The two sentences have more or less the same meaning. Yet in (5.18), it is the action (*t-talang* 'CV.FUT-run'), rather than its reason or beneficiary (*hii=mu* 'my body'), that is in the reduplicated (Ce-STEM) form. Conversely, the reduplicated form in (5.19) (*k-k-brax hii=mu* 'my body's being strong') expresses the reason for an action (*t<m>alang=ku* 'I run'). Due to these functional and semantic differences, the two usages of reduplication must be distinguished.

5.4.2.3. Instrument nominalization

No Ce- reduplicated form suffixed with *-an* is attested in Tsukida. As discussed in §3.6.5, there exist, in fact, such forms e.g., *m-mah-an* 'drinking container' > *imah* 'to drink'. Note that these forms are unequivocally nominal and refer to instruments. This pattern of *Ce*-STEM-*an* seems to resonate with Blust's (1998) claim that PAn *Ca- reduplication, *with or without* the additional *-an suffix, formed instrument nominals. Yet in Truku, this morphological process is synchronically unproductive. Novel combinations are not accepted by native speakers e.g., **t-cinun-an* '(intended) loom/weaving instrument'[68] > *cinun* 'to weave'. Furthermore, *Ce-…-an* forms do not name specific items. For instance, *m-mah-an* does not mean 'cup' but is rather translated as 'a drinking tool'. The concept of a 'cup' in Truku is expressed as *kopu* (probably a loan from Japanese *koppu*, itself an adaptation of English *cup*) or *kopu m-mah-an* (lit. 'drinking cup'). Neither does *t-tkan-an payay* (*tkan* 'to pound' + *payay* 'unhulled rice') mean 'pestle' or 'mortar'. Instead, it is better translated as 'tool for pounding rice' as in (5.20).

[68] Truku has an underived nominal for 'loom' *ubung*, and the general concept of 'weaving instruments' is expressed as *qngqaya t-cinun* ('tool RED-weave').

(5.20) sruw o, t-tkan-an payay
pestle TOP INST.NMLZ-pound-NMLZ unhulled.rice.OBL
'A pestle is a tool for pounding rice'.

5.4.2.4. Means/manner nominalization

In contrast, Ce- reduplication is highly productive in means/manner nominalization (§3.6.6). Ce- reduplicated forms refer either to means for the action denoted by the verb or manners in which an action is performed. Although Tsukida treats these instances as a "nominal usage" (headless relative clause) of CV future forms, I consider them a distinct morphological process, since it diverges semantically from CV. Specifically, reduplicated forms like those found in (5.21–22) do not refer to an instrument, beneficiary, theme, or reason.

(5.21) ini=ku skuxul ka t-tgsa=na
NEG=1SG.NOM like.AV.NFIN NOM MAN.NMLZ-teach=3SG.GEN
'I do not like his way of teaching.'

(5.22) yayu o, k-krut damat
kitchen.knife TOP MAN.NMLZ-cut vegetable.OBL
'A kitchen knife is for cutting vegetables.'

By the same token, means/manner nominalization is also differentiated from subjunctive gerunds on semantic grounds, since the latter refer to events.

Like *Ce-…-an* instrumental nominals, most Ce- reduplicated nominalizations in Truku generally do *not* refer to actual named objects. The online dictionary of Truku lists just a few that do. The translation in the dictionary was given in Chinese.

(5.23) a. *ciyu* 'to point' : *c-ciyu* 指標 'sign, signpost'[69]
b. *csiyus* 'to stir fry' : *c-csiyus* 食油 'cooking oil'
c. *hilaw* 'to cover' : *h-hilaw* 被子'blanket'
d. *tgsa* 'to teach' : *t-tgsa* 教材 'teaching material'

In addition, a language consultant claims that *p-patas* ('RED-write') refers to a pen or pencil (i.e., a writing instrument). However, an older consultant informed me that the more traditional form for 'pen/pencil' is *buji p-patas* ('arrow RED-write'), literally 'arrow(-like object) for writing'. Thus, it is possible that *buji* 'arrow' got dropped to give rise to the truncated, fossilized form *ppatas* 'pen'. I have identified other instances of a noun followed by a Ce-reduplicated form (5.24). Each of these can be treated either as a manner/means nominalization or a CV future form

[69] Retrieved on 11/28/16 from 原住民族語言線上詞典 (http://e-dictionary.apc.gov.tw/trv).

modifying a noun with its intended purpose. Note that the noun in these fixed phrases cannot be omitted.

(5.24) a. qumi s-sais
needle RED-sew
'sewing needle'
b. qumi b-beytaq m-narux seediq
needle RED-stab STAT.FIN-sick person
'injection needle' (lit. needle for stabbing sick people)
c. qngqaya t-cinun
things RED-weave
'weaving tools'

This situation is in stark contrast to the Ca- reduplication in languages like Saisiyat and Thao, which creates terms for named objects.

(5.25) *Saisiyat* (Yeh 2011:566)
a. ha-hiyop ka bokes
NMLZ-blow ACC hair
'hair dryer'
b. pa-pori'
NMLZ-apply (balms)
'plaster'
c. pa-potoy
NMLZ-wrap
'bandage'

(5.26) *Thao* (Blust 1998:54–55)
a. /capu/: *c<m>apu* 'sweep': *ca-capu* 'broom'
b. /cput/: *c<um>put* 'filter, strain': *ca-cput* 'seive, strainer'
c. /kishkish/: *k<m>ishkish* 'shave': *ka-kishkish* 'razor'
d. /runrun/: *m-runrun* 'roll something up': *ra-runrun* 'a bandage'

In fact, it is not clear whether Truku has a productive morphological process for deriving instrument nominals. There are a number of verb roots that are homophonous with the tool used for the actions they stand for (table 5.3). Although it is clear that one is derived from the other via conversion, the directionality is indiscernible based on synchronic, language-internal grounds. It is equally possible to analyze the verbs as having derived from the nouns as it is to analyze the nouns to have derived from the verbs.

Table 5.3. Homophonous verb root and instrumental nominal pairs in Truku Seediq

verbal root	instrument nominal
gumuk 'to cover'	*gumuk* 'lid'
huqut 'to walk with a cane, to use sth as a cane'	*huqut* 'cane'
kalu 'to comb (hair)'	*kalu* 'comb'
krut 'to cut'	*krut* 'saw'
sbut 'to hit'	*sbut* 'whip/cane'
sksik 'to sweep'	*sksik* 'broom'
ciyux 'to comb (hair)'	*ciyux* 'comb'
tpnu 'to rest one's head on a pillow, to use sth as a pillow'	*tpnu* 'pillow'
tucing 'to fall, to hammer sth'	*tucing* 'hammer'

To recapitulate, Ce- reduplication in Truku productively derives CV future verb forms, subjunctive gerunds, and means/manner nominalizations. There are some instrumental nominals which contain the Ce- reduplicant (often in conjunction with the suffix *-an*), but the productivity of the process is limited.

5.4.3. Ce- reduplication + affix other than *-an*

Besides *-an*, Ce- may co-occur with other affixes for various purposes: (i) with *p-*, to mark reciprocity, (ii) with *s-*, to derive a verb 'to smell of [noun]' from a noun, and (iii) with *t-*, to derive a verb 'to be habitually in search of [noun]'. Tsukida lists another verb-deriving function for which only one example has been found (*pk-i-ayug* 'to be wrinkly' supposedly from *ayug* 'valley'). Since it is not possible to surmise a pattern based on one example, this will not be discussed below.

5.4.3.1. Reciprocal marking

Ce- reduplication with the prefix *p-* (p-Ce-STEM) creates the reciprocal form of a verb e.g., *p-s-sikul* 'to push each other', *p-t-tlung* 'to touch each other'. In the presence of the finite AV marker *m-*, *p-* is deleted due to pseudo nasal substitution[70]. As discussed earlier, Zeitoun and Wu (2006) conclude that this pattern is in fact Ca- reduplication, since reciprocal verb forms are also created via Ca- reduplication in Atayal, Paiwan, Puyuma, Maga Rukai, Mantauran Rukai, and Saisiyat.

[70] According to Blust (2004), pseudo nasal substitution is a morphonological process found across AN languages that targets roots with an initial bilabial /p/ or /b/. Upon affixation of a reflex of PAn *<um> (AV), either the initial consonant or the entire initial syllable is deleted e.g., Truku *patas* > *m-atas* 'AV-write', *barig* > *m-arig* 'AV-buy'. Thus, it creates an illusion as if /m/ substituted for the root-initial consonant.

(5.27) m-d-dayaw ka Rubiq ni Masaw
AV-RED-help NOM Rubiq and Masaw
'Rubiq and Masaw help each other.' (Tsukida 2009:672)

(5.28) s<m>nru ka rudan, p-d-dayaw
preach<AV> NOM eldery RECP-RED-help
'Our elders preach (us), "help one another".' (field note)

5.4.3.2. Verbal derivation: '*Smell of X'*

The combination of the *s-* prefix and Ce- reduplication on a noun yields a verb in the form of s-Ce-NOUN with the meaning 'to smell of X'. The *s-* prefix is a verbalizer used to derive verbs from nouns, though the semantics of the resulting verbs is idiosyncratic e.g., *dara* 'blood': *s-dara* 'to bleed', *bgihur* 'wind': *s-bgihur* 'to blow (of a wind)', *quwaq* 'mouth': *s-quwaq* 'to be noisy, mouthy'.

(5.29) 'smell of X' (Tsukida 2009:268)
a. *blbul* 'banana': *s-b-blbul* 'to smell of a banana'
b. *mami* 'orange' : *s-m-mami* 'to smell of an orange'

(5.30) s-q-qsurux bi ka ruwan sapah nii
VBLZ-RED-fish INT NOM inside house PROX
'This room smells like fish.' (field note)

In comparison, the format s-Cu-NOUN in Tgdaya Seediq has a more general meaning of 'similar to X'.

(5.31) *Tgdaya Seediq* (Holmer 1996:56)
a. *rodux* 'chicken' : *s-r-rodux* 'tasting like chicken'
b. *kuung* 'darkness, dark' : *s-k-kuung* 'feeling dark'
c. *ngiro* 'mushroom' : *s-ng-ngiro* 'tasting like mushrooms'

5.4.3. Verbal derivation: '*Always in search of X'*

The prefix *t-* in conjunction with Ce- reduplication on a noun forms t-Ce-NOUN, meaning 'to be always in search of X'. Like *s-*, *t-* also has a verbalizing function e.g., *ngahan* 'name': *t-ngahan* 'to name', *laqi* 'child': *t-laqi* 'to give birth', *hidaw* 'sun': *t-hidaw* 'to be exposed to sunlight, to expose sth to sunlight' (Tsukida 2009:262–263).

(5.32) 'to be always in search for X' (Tsukida 2009:266)
a. *uwa* 'young woman' : *t-e-uwa* 'to always pursue young women'
b. *pila* 'money' : *t-p-pila* 'to always pursue money'

In the sections above, I provided a description of Truku reduplication, which has three formal variants: C_1e-/C_1eC_2e-, Ce-, and Ce- + affix. C_1e-/C_1eC_2e- marks plurality on nouns and adjectives. Ce- performs functions of CV future verb formation, subjunctive gerund formation,

means/manner nominalization, and occasional instrumental nominalization. Lastly, Ce- acts in conjunction with other affixes to derive reciprocal verb forms as well as verbs 'to smell of [noun], and 'to be always in search of [noun]'.

The next section will reveal that future tense marking is a function of Ca-(-like) reduplication across Atayalic subgroup of the Austronesian language family. Unlike Truku, some Atayal varieties such as Mayrinax (L. Huang 2002) also use the morphological process productively for the formation of instrument nominals. Although formation of means/manner nominalization via Ca- reduplication is not widely attested, I have evidence to suggest that it at least exists in Tgdaya Seediq. L. Huang (2002) also makes a passing remark on the pattern for Mayrinax.

5.5. Ca- reduplication in other Atayalic varieties

5.5.1. C(u)- reduplication in Tgdaya Seediq

Holmer's (2002) paradigm for the Tgdaya variety of Seediq shows that reduplication is found on PV- and LV-marked forms to indicate both immediate and distant future (cf. figure 4.3). The pattern of reduplication is described as "reduplication of the initial consonant" in Holmer (2002: 336) and as CV- in Holmer 1996. Though reduplicated forms are phonetically [Cu-]STEM, the aforementioned pre-penultimate vowel reduction rule obscures the underlying phonological value of the reduplicant. Typically, [u] in Tgdaya is not written where its presence is predictable from the context. The same writing convention applies to the schwa in Truku.

However, he states that "the forms listed as [future] for [LV] and [PV] are very specialised in meaning…and are certainly not prototypical future forms" (Holmer 2002:334). Thus, reduplication is most likely not highly productive in the verbal paradigm of the dialect. Regarding the use of reduplication with the PV suffix *-un*, Holmer states that it "emphasises the truth value of the proposition" (2002:336), as exemplified by the difference between (5.33a) and (5.33b).

(5.33) *Tgdaya Seediq* (Holmer 2002:336)

a. biq-un=su=mu
   give-PV=2SG.NOM=1SG.GEN
   'I will give it to you.'

b. **b**-biq-un=su=mu!
   RED-give-PV=2SG.NOM=1SG.GEN
   'I *will* give it to you.'

Also, Holmer states that a reduplicated form with the LV marker *-an* "is entirely nominal in meaning and refers to the possibility of an action occurring. It is usually used together with the verbs *niqan* 'there is' and *uka* 'there is not'" (2002:337).

(5.34) Tgdaya Seediq (Holmer 1996:442)

a. uka **m**-mah-an sino mukan
NEG.EXIST RED-drink-LV wine Taiwanese
'Taiwanese wine is undrinkable.' (lit. 'There is no drinking Taiwanese wine.')

b. niqan **u**-uq-an damac kiya
EXIST RED-eat-LV food that
'This food can be eaten (although it is not very good).'

Note that in this usage, *-an* seems to lack the LV function despite being glossed as such. Holmer further maintains that these patterns of reduplication in Tgdaya do not necessarily indicate futurity, but rather have to do with an action being contemplated: "…both of these forms refer to an action which has not taken place and is not taking place at the time of utterance, but which is being contemplated as a definitive ([PV]) or hypothetical ([LV]) future action" (1996:43). In fact, Holmer admits to treating them as future tense forms merely for the sake of symmetry in the verbal paradigm, thereby providing PV and LV forms with parallels with the AV future form.

In fact, this combination of reduplication and PV or LV marker is found in Atayal (cf. §5.4.4.2). Wolff 1973 reconstructs PAn *RED…-en and *RED…-an as future/general action forms in PV and LV, respectively, based on evidence from Atayal and Samar-Leyte (Visayan language of the Philippines). On the other hand, corresponding AV and CV forms are missing from his PAn paradigm. On the two reduplicated forms, Wolff states as follows:

> *This category is part of the verbal inflectional system in [Atayal] and [Samar-Leyte]. In other Austronesian languages the category survives as a productive type of nominal derivation or as an unproductive or petrified formation. It is not certain what the meaning of this form was in the protolanguage and whether it was really part of the verbal inflectional system or simply a productive type of nominal derivation which has become part of the verbal inflectional system independently in [Samar-Leyte] and [Atayal]" (Wolff 1973:88).*

Therefore, the patterns seem to be reconstructable at least at the Proto-Atayalic level. Whether they predate Proto-Atayalic, as well as their actual function, is uncertain.

According to Holmer, Tgdaya also has a small number of AV forms that are reduplicated. The base for reduplication itself is in the future tense form e.g., ***m-mosa***[71] 'RED-go.FUT' and ***m-maha***

[71] Listed as *musa* in Holmer 1996, but more accurately *mosa* /mosa/.

'RED-get.going.FUT' (1996:217). Furthermore, he suggests that many *-un* and *-an* suffixed forms *appear* to be non-reduplicated only due to a morphonological rule that bans the reduplicant to surface phonetically: "The C-prefix is only realised if it appears as the onset of the prestress [i.e., antepenultimate] syllable" (1996:216). Therefore, ***b**-biq-un* [bu'biɘqon] 'RED-give-PV', ***m**-mah-un* [mu'maχon] 'RED-drink-PV', and ***u**-uq-un* [o'oqon] 'RED-eat-PV' are attested, but not ****d**-dndang-un* [dudunu'daŋun] 'RED-boil-PV' or ****b**-brig-un* [bubu'ligun] 'RED-buy-PV'. On the contrary, I have recorded a number of forms in which a reduplicant preceding the antepenult is pronounced e.g., ***s**-slahay-un* [susuɬa'χayun] 'RED-study-PV' and ***d**-dehek-an* [dede'χekan] 'RED-arrive-LV'.

I have found a few additional examples of reduplication with *-un* or *-an* in my small corpus of Tgdaya data. Many of them appear to be used nominally, being preceded by the nominative case marker *ka*. The number of samples is too small to discern the difference between RED-...-*un* forms and RED-…-*an* forms. Nonetheless, (5.38) seems to confirm Holmer's claim that *-an* suffixed forms stand for an action (or event, perhaps), whereas the *-un* suffixed form in (5.39) evidently refers to the *object* of learning, rather than the action of learning itself.

(5.35) uxe beyo dehuk ka **s**-sa-un=mu di
NEG long.time arrive.AV NOM RED-go-PV=1SG.GEN CS
'My destination is almost there.'

(5.36) maanu ka **p**-paq-un=su
what NOM RED-hit-PV=2SG.GEN
'Which one are you going to kill?' (lit. 'That which you are going to kill is what.')

(5.37) b-bliq-un=mu **p**-puq-un
RED-make.good-PV=1SG.GEN RED-eat-PV
ka bobo m-n-narux=ku
NOM after STAT.FIN-PFV-sick=1SG.NOM
'Having been sick, I will eat better.' (*p-puq-un* = 'what to eat; diet')

(5.38) s-pru-un=na m-ita ka **s**-slahay-an=mu
CAUS-big-PV=3SG.GEN AV-see NOM RED-study-LV=1SG.GEN
'She considers my learning important.'

(5.39) egu riyung ka **s**-slahay-un=mu
a.lot INT NOM RED-study-PV=1SG.GEN
'I have a lot of things to learn.' (lit. 'That which for me to learn is many.')

(5.40) daha ali **d**-dehek-an hini=na ka Rabe
two day RED-arrive-LV here=3SG.GEN NOM Rabe
'It will take Rabe two days to get here.'

Although unmentioned by Holmer, Tgdaya also seems to have CV future forms: Cu-reduplicated forms with no additional affixes. Just as in Truku, CV future forms choose the

theme (5.41–42) and reason (5.43) as the pivot. I am yet to encounter instances where an instrument or a beneficiary serves as pivot, but I expect them to exist.

(5.41) ima ka **a**-adis[72]=su hini
who NOM RED-bring=2SG.GEN here
'Who are you bringing here?'

(5.42) **a**-adis na Rabe ka brkawe nii
RED-bring GEN Rabe NOM plum PROX
'These plums are for Rabe to bring.'

(5.43) maanu ka **s**-sange=su ali saya
what NOM RED-rest=2SG.GEN day now
'What are you resting for today?'

There are also a few examples that resemble means/manner nominalization in Truku.

(5.44) naqah **q**-qilit=mu ringo
bad.STAT.FIN RED-peel=1SG.GEN apple.OBL
'I am bad at peeling apples.' (lit. 'My way of peeling apples is bad.')

(5.45) m-sliko ba **t**-talang na Rabe
STAT.FIN-fast INT RED-run GEN Rabe
'Rabe runs fast.' (lit. 'Rabe's manner of running is fast.')

(5.46) so nii ka **h**-huqil na quyuw
SIM PROX NOM RED-die GEN snake
'Snakes die like this (lit. 'Snakes' way of dying is like this.')

There is some evidence for gerundive usages of Cu- reduplication equivalent to those of Ce- reduplication in Truku. In (5.47–48), reduplicated forms assume the sentential pivot position and refer to events that are yet to be realized.

(5.47) betaq knuan ka **g**-gduru=su
until when NOM RED-rest=2SG.GEN
'Until when are you going to rest?'

(5.48) betaq knuan ka **l**-lingis=su
until when NOM RED-cry=2SG.GEN
'Until when are you going to cry?'

[72] While the realization of the reduplicated vowel here as [a] might seem to lend support to the CV- (as opposed to Ca-) analysis, this is most likely due to a morphonological process that assimilates the antepenultimate vowel to the penultimate one across /h/ or when there is no intervening consonant (e.g., *mi-icu* 'STAT.FIN-be scared' /m**u** + iʈsu/ [m**i**'iʈsu], *slahay-an* 'study-LV' /sul**u**hay + an/ [sul**a**'χayan]).

Compare:

(5.49) *Truku Seediq*

mp-beytaq	knuan	ka	**k**-k-saang=su	hug?
AV.IRR-until	when	NOM	SBJV.GER-STAT.NFIN-angry=2SG.GEN	Q

'Until when are you going to be angry?'

There is also an example of a Cu- reduplicated form in the predicate position that indicates a purpose (5.50). This is comparable with the Truku pattern exemplified in (5.51).

(5.50)

k-k-malu=su	maanu	ka	so	nii=su	
RED-STAT.NFIN-good=2SG.GEN	what	NOM	SIM	PROX=2SG.GEN	

'What good comes from you being like this?' (lit. 'You are like this for your being good in what [way]?')

(5.51) *Truku Seediq*

k-k-brax	hii=mu	ka	t<m>alang=ku
SBJV.GER-STAT.NFIN-strong	body=1SG.GEN	NOM	run<AV>=1SG.NOM

'I run for my health.' (lit. '(The reason why) I run is for my body's being strong.')

Despite these parallels, subjunctive gerunds do not have completely overlapping distribution between Truku and Tgdaya. The following are attempts to elicit Cu- reduplicated forms in the environment after a verb of causation, speech act, or expectation. All but one was rejected by native speakers. Cases like (5.53) and (5.54) show that a full clause is preferred as a complement of such verbs.

(5.52)

*ga=ku	t<um>ara	**a**-adis=su	damac	
PROG=1SG.NOM	wait<AV>	RED-bring=3SG.GEN	dish.OBL	

'I am waiting for you to bring food.'

(5.53) a.

psmyah-un=mu	ba	m-phepah	ka	boxin=mu
expect-PV=1SG.GEN	INT	AV-bloom	NOM	lily=1SG.GEN

b.

*psmyah-un=mu	ba	**p**-phepah	boxin=mu
expect-PV=1SG.GEN	INT	RED-bloom	lily.GEN=1SG.GEN

'I am looking forward to my lilies' blooming.'

(5.54) a.

mu-durun=ku	Rabe
AV.IRR-request=1SG.NOM	Rabe.OBL
mu-tabu=Ø	rudux=mu
AV.IRR-feed=3SG.NOM	chicken.OBL=1SG.GEN

b.

?mu-durun=ku	Rabe
AV.IRR-request=1SG.NOM	Rabe.OBL
t-tabu=na	rudux=mu
RED-feed=3SG.GEN	chicken.OBL=1SG.GEN

'I will ask Rabe to feed my chickens.'

(5.55) a. mu-durun=ku Rabe ma-adis=Ø brkawe
AV.IRR-request=1SG.NOM Rabe.OBL AV.IRR-bring=3SG.NOM plum.OBL
b. mu-durun=ku Rabe **a**-adis=na brkawe
AV.IRR-request=1SG.NOM Rabe.OBL RED-bring=3SG.GEN plum.OBL
'I will ask Rabe to bring plums.'

In conclusion, Cu- reduplication in Tgdaya shares many, if not all, functions of Truku Ce- reduplication. It indicates futurity or irrealis mood when combined with a voice marker (AV irrealis *mu-*, PV *-un*, LV *-an*, and CV *Ø*), and it is used for means/manner nominalization. Some subjunctive gerund-like usages have been documented, although their level of productivity is yet to be determined. The following section will explore the functions of Ca- reduplication in Atayal, the other member of the Atayalic subgroup. The discussion will focus on Mayrinax Atayal due to the process being phonologically transparent in this variety, as well as its analysis in the literature being more refined compared to that of other dialects.

5.5.2. Ca- Reduplication in Mayrinax Atayal

Atayal has a wide geographical spread, spoken in Taipei, Taoyuan, Hsinchu, Miaoli, Taichung, Nantou, Hualien, and Ilan Counties of Taiwan. The language is split into two major dialect groups: Squliq and C'uli'. According to Li (1981:235), Squliq is "fairly uniform" while C'uli' varieties "can be quite divergent from each other." Mayrinax, a C'uli' variety, is said to be the most conservative of all. One phonological difference among Atayal varieties that is of particular interest in the present context is the presence/absence of the vowel reduction rule. Recall that in Seediq varieties, most pre-penultimate vowels reduce to the default quality: [u] in Tgdaya and [ə] in Truku (also purportedly in Toda (Lee 2012), although not much is known about this variety). Squliq Atayal also largely limits pre-penultimate vowels to the schwa (Rau 1992). For instance, the vowel phonemes /i/ and /u/ in the penult of the roots *kita* 'see' and *huziq* 'wet', respectively, are both realized as the schwa when they move to the antepenult as a result of suffixation.

(5.56) *Squliq Atayal* (H.-C. Huang 2006:491)
a. *ktan* /k**i**ta + an/ [k**ə**tan] 'see-LV'
b. *hziqan* /h**u**ziq + an/ [h**ə**ʒiqan] 'wet-LV'

Discordantly, it appears that at least some C'uli' varieties including Mayrinax have not developed this rule.

Distribution of the schwa is said to be contextually predictable in Squliq, and it never occurs in a stressed syllable (Li 1980, Rau 1992). Unless it appears due to reduction or in a transition

between a high vowel and a /q/ or /h/, Li treats it as an epenthetic vowel that is inserted to break up underlying consonant clusters. Thus, /hpah/ 'flower' is realized as [həpæh], /bbuʔ/ 'jungle, bush' as [ßəßuʔ], and /qhniq/ 'bird' as [qəhəniəq] (Li 1980:355). This results in a situation that Li describes as follows: "[there] is little limitation on the number or order in the combination of consonants before the penultimate vowel" (1980:357). It is not clear why the schwa in these examples cannot be posited in the underlying representation to avoid such a peculiarity.

Perhaps due to this treatment of the schwa, the reduplicative pattern found in Squliq Atayal is termed C- reduplication by Zeitoun and Wu (1996), whereby only the first consonant is copied. Subsequently, automatic schwa epenthesis breaks up the resulting initial consonant cluster. Thus, the result is phonetically [Cə]. Just as in Truku, any leftward reduplication (regardless of the underlying form, which could be CV-, Ca-, Cə-, or C-) would theoretically result in the same surface form. Nevertheless, Lin (2015) points out that the reduplicant of not only di- and tri-syllabic roots, but even monosyllabic roots in Squliq is always the schwa. If the nucleus of the reduplicant is in the penultimate position, it would not undergo vowel reduction. Therefore, this essentially rules out both Ca- and CV- reduplication, which would yield **za-zik* or **zi-zik* instead of the actually attested form *zə-zik* 'very deep' (< *zik* 'below'), and **mə-ka-kut* or **mə-ku-kut* instead of *mə-kə-kut* 'to cut each other' (< *mə-kut* 'to cut') (examples from Zeitoun and Wu 2006:105). Although Lin raises the possibility that fixed vowel reduplication is involved (as [ə] as the fixed vowel instead of [a]), it is concluded "based on language internal evidence" (2015:83) that only the consonant is copied and that the schwa is epenthetic.

If the reduplicant indeed consisted of a consonant alone, one may question what happens when the stem to be reduplicated lacked an initial consonant. We observed earlier that in most languages CV- reduplication applied to a vowel-initial stem simply results in the vowel being copied as a reduplicant. Similarly, Ca- reduplication simply prefixes the vowel /a/ onto a vowel-initial stem. Unfortunately, it is not clear how the reduplicative process is applied in Squliq, since the analysis is dependent on the phonemic status of the word-initial glottal stop. Lin (2015) argues that the syllable in Squliq is minimally CV; that is, an onset is required. I believe that this claim is based on Li's (1980, 1981) treatment of the word-initial prevocalic glottal stop as a phoneme. To the contrary, Egerod (1966) and Rau (1992) consider the consonant to be non-contrastive in this position, since its presence is predictable. Under this analysis, the minimum

syllable requirement in Squliq would be V. Indeed, vowel-initial roots such as *a:ras*[73] 'to take along' and *u:sa* 'to go' are found in Rau (1992:27–28). Rau admits to having insufficient data to determine reduplication patterns for such roots.

This is where Mayrinax data proves to be particularly illuminating. Unlike Squliq and Seediq dialects, Mayrinax does not have a pre-penultimate vowel quality restriction. According to Lillian Huang (2000), Ca- reduplication in Mayrinax has two functions: irrealis/future marking and nominalization. Table 5.4 summarizes the verbal paradigm of Mayrinax Atayal.

Table 5.4. Mayrinax Atayal verbal paradigm (adapted from Huang 2000)

		AV	PV	LV	CV
realis	perfective	m-/ma-/<um>/ø	<in>	-an	si-
	imperfective		-un		
irrealis (future)		pa-	Ca-…-un	Ca-…-an	Ca-

Irrealis/future marking via Ca- reduplication applies to all NAV forms but not to AV forms. AV irrealis verbs take the form of *pa*-STEM. According to Huang's (2002) analysis, *pa*- indicates irrealis/future and AV is zero-marked.

(5.57) *Mayrinax Atayal* (Huang 2000:117)
pa-Ø-qaniq=ci' cu' bunga'
IRR-AV-eat=1SG.NOM ACC yam
'I will eat yams.'

As for irrealis/future NAV forms, PV and LV are marked by the same voice affixes as they are in the realis mood: *-un* and *–an*, respectively. On the other hand, CV is zero-marked. Thus, the format for PV, LV, and CV future/irrealis verbs is *Ca*-STEM-*un* (5.58a), *Ca*-STEM-*an* (5.58b), and *Ca*-STEM (5.58c–d), respectively.

(5.58) *Mayrinax Atayal* (Huang 2000:117–118)
a. na-niq-un nku' nabakis ku' bunga'
IRR-eat-PV GEN elderly NOM yam
'An/this old person will eat (these) yams.'
b. ha-hihip-an ni' Yumin 'i' Limuy
IRR-kiss-LV GEN Yumin NOM Limuy
'Yumin will kiss Limuy.'
c. ba-baiq-Ø nku' nabakis 'i' Limuy ku' xuil
IRR-give-CV GEN elderly DAT Limuy NOM dog
'An/this old person will give a/this dog to Limuy.'

[73] Although Rau marks vowel length, it is predictable and is thus not contrastive.

d. qa-quwah-Ø=cu ni' Tali' cu' salaman
IRR-wash.dishes-CV=1SG.NOM GEN Tali' ACC bowl
'Tali' will wash dishes for me.'

In addition to marking irrealis mood, Ca- reduplication in Mayrinax also forms argument nominals, which encode the semantic role of verbal arguments. Huang notes that each form resembles that of corresponding irrealis verbal form; agentive nominals resemble AV verb forms, patient nominals resemble PV verb forms, and so on. In fact, the two sets of forms are phonetically identical. Huang (2002) distinguishes them by consistently marking the morpheme boundaries in verbal forms but not in nominalizations. As discussed in §4.1, difficulty in distinguishing verbal forms from argument nominals is shared by Truku as well as many other Formosan languages.

Agent nominals are created with either *pa-* (5.59) or *m-/ma-/<um>* (5.60), which are homophonous with the irrealis/future marker and the realis AV marker, respectively.

(5.59) *Mayrinax Atayal* (Huang 2002:211–212)
a. *paquwas* 'to sing': ***pa****paquwas* 'singer, one who will sing'
b. *quriq* 'to steal': ***pa****quriq* 'thief, one who will steal'

(5.60) *Mayrinax Atayal* (Huang 2002:211–212)
a. *paquwas* 'to sing' : ***m****aquwas* 'singer, one who is singing'
b. *quriq* 'to steal': *q****um****riq* 'thief, one who is stealing'

Patient nominals in the irrealis mood are formed with *Ca-…-un* (5.61), whereas those in the realis mood are formed with <*in*> (5.62). In my description of nominalization in Truku (Chapter 3), I treated the corresponding (*Ce-*…)-*un* marked forms and <*n*> marked forms as patient nominalization and result nominalization, respectively.

(5.61) *Mayrinax Atayal* (Huang 2002:215)
a. *qaniq* 'to eat' : ***na****niq****un*** 'food, things to be eaten'
b. *patauwaw* ' to work': ***pa****pwa****un*** 'work to be done'

(5.62) *Mayrinax Atayal* (Huang 2002:215)
a. *patiq* 'to write': *p****in****atiq* 'word, handwriting'
b. *tahuq* 'to cook': *t****in****ahuq* 'cooked dish'

Some patient nominals are also formed with <*in*>*…-an*.

(5.63) *Mayrinax Atayal* (Huang 2002:217)
a. *timami* 'to pickle': *t****in****mami****an*** 'pickled meat'
b. *ciyok* 'to answer': *t****in****yok****an*** 'answer'

Locative nominals are constructed either via *Ca-…-an* (irrealis) or *<in>…-an* (realis).

(5.64) *Mayrinax Atayal* (Huang 2002:217–219)
a. *qaniq* ‘to eat’: ***na****niq****an*** ‘restaurant, table, place one will eat’
b. *thawnaq* ‘to sit’: ***ta****thawnaq****an*** ‘chair’
c. *ɣhapuy* ‘to cook’: ***ɣa****ɣhapuy****an*** ‘kitchen, cooker, place one will cook’

(5.65) *Mayrinax Atayal* (Huang 2002:217–219)
a. *ɣhapuy* ‘to cook’: *ɣ****in****hapuy****an*** ‘place where one cooked’
b. *kat* ‘to bite’: *k****in****ac****an*** ‘place being bitten’

Like CV irrealis forms, instrumental nominals take the form of *Ca*-STEM. Unlike Ce-reduplicated forms in Truku, many of these items evidently refer to named objects.

(5.66) *Mayrinax Atayal* (Huang 2002: 219)
a. *patiq* ‘to write’: ***pa****patiq* ‘pen’
b. *hawßin* ‘to cut’: ***ha****hawßin* ‘instrument used to cut’
c. *kaloʔ* ‘to comb’ ***ka****kaloʔ* ‘comb’
d. *kiri* ‘to grind’: ***ka****kiri* ‘knife grinder’
e. *capux* ‘to sweep’: ***ca****capux* ‘broom’
f. *hirhir* ‘to saw’: ***ha****hirhir* ‘saw’
g. *samaɣ* ‘to place (mat)’: ***sa****samaɣ* ‘mat’

In a footnote, Huang (2002) maintains that she “also collected sentences like the following…Whether the underlined nominals below refer to actions or instruments (thus the forms are irrealis [CV]) deserves further investigation.”

(5.67) *Mayrinax Atayal* (Huang 2002:209 footnote, glosses mine)

Ø-ßalaiq	aʔ	**qa**qaniq=niaʔ,	
AV-good	NOM.NRF	eating=3SG.GEN	
ruʔ	Ø-ßalaiq	aʔ	**qa**qilaap=niaʔ
and	AV-good	NOM.NRF	sleeping=3SG.GEN

‘Lit. His eating is good, and his sleeping is good.’
‘What he eats is good, and the place he sleeps in is good.’

Though Huang’s translation may suggest differently (the reduplicated forms translated as an object and a location: “what he eats” and “the place he sleeps,” respectively), the constructions are reminiscent of the manner/means nominalization in Truku, whereby the reduplicated forms would be translated as “the way he eats/sleeps.”

To recapitulate, the analysis of reduplication in Squliq requires further refinement both at phonological and functional levels. On the other hand, data from Mayrinax Atayal indicate that the dialect has Ca-, but not CV-, as a reduplicant. Mayrinax Ca- reduplication, like Ce-reduplication in Truku and Cu- reduplication in Tgdaya, marks future tense/irrealis mood. There

appears to be one difference: the process does not apply to AV forms in Mayrinax, while it does so, albeit marginally, in Seediq dialects. Reduplication is also found on patient, locative, and instrumental nominals.

5.6. Proposed historical change of Ca- reduplication leading to Truku subjunctive gerunds

Comparison among Atayalic varieties proves that irrealis/future marking is a shared function of reduplication, though its productivity varies. Due to the productive vowel reduction rule in most varieties, the form of the reduplicant may be analyzed variously as Ca-, CV-, Ce([Cə])-, Cu-, or C- in synchronic terms. However, based on evidence from Mayrinax Atayal, we can infer that the reduplicant for future/irrealis marking was *Ca- in Proto-Atayalic, the common ancestor of Atayal and Seediq. On the assumption that the formation of instrumental nominal via Ca-reduplication can be reconstructed to PAn per Blust (1998), Mayrinax Atayal reflects it as Ca-, whereas the function seems to have been largely lost in Seediq varieties. This would imply that the two functions of *Ca- existed side by side in Proto-Atayalic.

In Seediq, future marking via reduplication is strictly optional. Particularly striking is the fact that when reduplication co-occurs with AV-marking, the verbal stem is already in the irrealis form. Thus, this function is merely emphatic. Since reduplication does not co-occur with AV markers in Mayrinax, it is not clear whether it did so in Proto-Atayalic. In both Mayrinax and Seediq, there is a split in the voice-marking patterns between PV and LV, on the one hand, and CV, on the other (table 5.5). Specifically, the former involve voice affixes shared between realis and irrealis forms, while the latter does not. CV is marked by the reflex of PAn *Si- in the realis mood, and is zero-marked in the irrealis mood.

Table 5.5. NAV irrealis/future morphology in Mayrinax Atayal, Tgdaya Seediq, and Truku Seediq

	PV	LV	CV
realis/finite	-un	-an	si- (Mayrinax) s- (Tgdaya, Truku)
irrealis/future	RED-…-un	RED-…-an	RED-

Because of the lack of morphological affinity between the two, I suspect that the realis CV and irrealis CV forms originated in two different sources: the former in the PAn Instrument Voice form and the latter in the instrumental nominalization via *Ca-. Some instrument nominals occurring in the predicate position could have been reanalyzed as verbal forms.

(5.68) Reanalysis of instrumental nominals as CV future forms

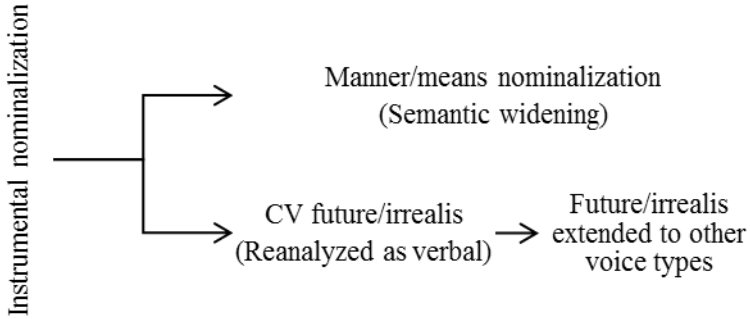
a. Ca$_{\text{INST.NMLZ}}$-VERB=1SG.GEN NOM NOUN ‘NOUN is my tool for VERB-ing’

reanalyzed as:

b. Ca$_{\text{CV.FUT}}$-VERB=1SG.GEN NOM NOM NOUN ‘NOUN is what I will VERB with’

Manner/means nominalization can also be viewed as stemming from instrumental nominalization, since an instrument is a means of performing an action. Instrument nominalization via reduplication ceased to be productive in Seediq varieties or, perhaps, became subsumed under manner/means nominalization, which obtained more general meanings due to semantic widening. These hypothetical processes of change are schematized in figure 5.1.

Figure 5.1. Proposed paths of change in functions of Ca- reduplication in Seediq



We saw that a few Formosan languages outside of the Atayalic branch (Amis, Puyuma, and Saisiyat) also have future/irrealis marking as one of the functions of Ca- reduplication. I reiterate that actual application of Ca- reduplication is far from being uniform across these languages (table 5.2). In Amis, the process appears to apply regularly to AV- and PV-marked forms, whereas its co-occurrence with LV and CV is not noted (Wu 2006). While Ca- may combine with all NAV markers (PV *-en*, LV *-an*, and CV *i-*) in Puyuma, neither Atayalic languages nor Saisiyat allows co-occurrence of future/irrealis C(a)- and an overt CV marker. Conversely, no overt AV marker is found on the Ca- reduplicated AV future verbs in Puyuma (Ca-STEM). Lastly, Saisiyat only allows future marking via Ca- reduplication on CV forms. Therefore, there is no strong evidence that these phenomena are historically related to one another.

How does the function of subjunctive gerund formation fit in this picture? Given that this function of nominalization is unattested outside of Atayalic, I conclude that it was a development after its split from the rest of Austronesian. Yet, its inception cannot be narrowed down any further, since its usage is not widely documented within Atayal. The origin of the subjunctive usage can be attributed to either means/manner nominalization or future/irrealis marking on

semantic grounds. Some instances of subjunctive gerunds like (5.69–70) can be paraphrased as '*How* my body becomes strong is by my running' or 'Your *way* of being angry is until when?'. However, the same cannot be said of cases in which subjunctive gerunds appear as a complement of another verb (5.71).

(5.69) k-k-brax hii=mu ka t<m>alang=ku
SBJV.GER-STAT.NFIN-strong body.GEN=1SG.GEN NOM run<AV>=1SG.NOM
'I run for my health.' (lit. '(The reason why) I run is for my body's being strong.')

(5.70) mp-beytaq knuan ka k-k-saang=su hug?
AV.IRR-until when NOM SBJV.GER-STAT.NFIN-angry=2SG.GEN Q
'Until when are you going to be angry?' (lit. 'Your being angry is until when?')

(5.71) d<m>udug knan g-geuguy=mu pila bubu=mu
urge<AV> 1SG.OBL SBJV.GER-steal=1SG.GEN money.OBL mother=1SG.GEN
ka Kuras
NOM Kuras
'Kuras urged me to steal my mother's money.' (lit. 'Kuras pushed me for my stealing my mother's money.')

Subjunctive gerunds are less likely to have been a direct extension from CV future marking, considering that the former are voice-neutral. The irrealis function must have neutralized over all (except AV, perhaps) voice types before the subjunctive usage emerged.

5.7. Conclusion

In this chapter, I proposed that the subjunctive gerund-deriving Ce- reduplication in Truku Seediq and Cu- reduplication in Tgdaya Seediq should be treated as a development of PAn *Ca- reduplication. Though historical change has rendered the vowel quality in Ce-/Cu- reduplication opaque, its continuity with the Ca- reduplication in Mayrinax Atayal strongly suggests that they were derived from Proto-Atayal *Ca- reduplication, which is in turn a reflex of PAn *Ca- reduplication.

Reduplication, in general, is known to exhibit a wide array of usages both cross-linguistically and language-internally. As far as Truku is concerned, a single pattern of Ce- reduplication was shown to productively derive three distinct constructions: means/manner nominalizations, future tense verb forms (especially in CV), and subjunctive gerunds. I argued that the function of subjunctive gerund-marking was an extension of either means/manner nominalization or future/irrealis marking. Data from Mayrinax Atayal suggest the possibility that the gerund-like usage of Ca- reduplication may be more widespread in Atayalic varieties, though a more detailed account is necessary to determine this point. I put forward a hypothesis in Chapter 4 that the

emergence of indicative gerund morphology involved the loss of thematic orientation from the voice marker/thematic nominalizer <*n*>, *-an*, and *-un*. The development of subjunctive gerund marking was accompanied by a similar process, such that the resulting constructions refer to events, rather than making a reference to thematic roles such as instrument, beneficiary, theme, reason, manner, or means.

Cases of Turku gerund morphology presented in Chapters 4 and 5 provide two examples of historical change that involves syntactic reanalysis originating in semantic ambiguity. This syntactic reanalysis, in turn, resulted in the functional extension of morphology. In both cases, a high level of polysemy ensued since the original functions of the morphemes were preserved. These findings challenge our understanding of the role of human cognition in processes of syntactic change as well as polysemy. To the extent that studies of cognition and historical change have heavily focused on lexical semantics (e.g., Blank and Koch 1999) and grammaticalization (e.g., Heine 1993, Heine, Claudi, and Hünnemeyer 1991), instances of syntactic reanalysis as these (in the sense of reanalysis without grammaticalization, per Haspelmath 1998) present another facet of exploration into mental linguistic representation.

CHAPTER 6

CONCLUSION

6.1. Summary of findings

This dissertation investigated the construction of gerunds in Truku Seediq, a Formosan language, from descriptive, typological, and historical perspectives. The purpose of this dissertation is (i) to provide the first detailed descriptive account of gerunds in Truku, and (ii) to trace back their morphological origins. Findings of this study are recapitulated below.

In Chapter 3, I first established the definition of gerunds within the scope of this dissertation based on existing literature. I defined gerunds as event- or state-denoting forms that demonstrate both prototypically nominal characteristics and prototypically verbal/sentential characteristics. They are derived via a productive morphosyntactic process applied to verbal stems. Actual characteristics associated with nouns and sentences vary from language to language. Languages also differ as to which particular characteristics, as well as how many of them, gerunds exhibit.

I demonstrated that there are two types of gerunds in Truku that differ in morphology, distribution, and semantics, but show largely analogous morphosyntactic behavior. Indicative gerunds are formed by <*n*>, *-an, -un*, and most often by a combination of <*n*>…*-an*, and refer to events that have already taken place or whose occurrence is imminent. Subjunctive gerunds are formed via Ce- reduplication, and refer to hypothetical or conceived events. Gerunds appear in the positions for arguments, adjuncts, predicates, and objects of prepositions. The major trait that sets subjunctive gerunds apart from their indicative counterpart is the former's co-occurrence with verbs of causation (e.g., 'urge', 'request', and 'forbid') and expectation (e.g., 'wait'). I illustrated the following traits in association with Truku gerunds:

(6.1) Verbal/sentential properties of Truku gerunds
a. may contain the perfective aspect marker <*n*>
b. require a subject
c. may take oblique-marked Patient as a complement
d. manner modification involves verb serialization
e. generally incompatible with demonstratives

(6.2) Nominal properties of Truku gerunds
a. distributional patterns
b. generally lack nominative case
c. the subject is marked genitive, just as possessors are
d. negated like nouns
e. compatible with the preposition *asaw* 'because of/for'

f. cannot form temporal clauses with *siida* ‘when’
g. interact with external possessor constructions (EPCs)

Chapter 4 and Chapter 5 sought to identify the origins of morphology that derive the indicative set and the subjunctive set of Truku gerunds, respectively. Chapter 4 presented a broad survey of gerunds in Formosan languages. Data from the Tgdaya variety of Seediq confirmed the patterns found in Truku Seediq, both in morphological and syntactic aspects. Aside from Tgdaya, constructions comparable to gerunds in Truku were only sparsely attested. Most importantly, in languages where their existence was confirmed or suspected, their morphology showed little to no affinity with that of Truku, suggesting an independent development. I argued that indicative gerund morphology was derived from voice markers/thematic nominalizers. This process may have been motivated by the ambiguity between entity-reading and event-reading of headless relative clauses.

Chapter 5 discussed reconstructed and contemporary functions of Austronesian Ca-reduplication. Focusing on evidence from Atayalic varieties, I drew a historical connection between the gerund-forming reduplicative pattern in Truku and the instrumental nominal-deriving *Ca- reduplication in Proto-Austronesian. Once again, only Tgdaya proved to have subjunctive gerund-like usages of reduplication parallel to Truku. I argued that a semantic extension from either future tense marking or manner/means nominalization allowed Ce-reduplication to acquire this function.

6.2. Implications and contributions

This study first makes contributions at a descriptive level. Truku Seediq is an understudied, endangered language. While Naomi Tsukida’s grammar provides a broad view of the language, this dissertation captures its finer workings in specific aspects of the morphosyntax. Applied to language pedagogy, this description can be used to aid language learners in disambiguating homophonous forms, and eventually to use them in appropriate contexts.

The implications of these findings extend beyond the understanding of this particular language. At a broader, typological level, gerunds in Truku both display a number of traits that validate previous description of gerunds and offer unique perspectives to the research of this construction. Particularly, the interplay of gerunds and external possession is a novel finding. External possession is defined as a phenomenon in which the possessor and the possessum do not form a constituent in surface syntax. The subject of a gerund in Truku demonstrably behaves

on par with a possessor of a noun phrase in this regard. This adds to the array of diagnostics for nominality in a variety of constructions where a language allows external possession.

Also revealed was a language-specific restriction on the formation of gerunds: absence of voice alternation. This limitation is likely due to the morphological identity between voice inflection and gerund marking. This homophony is a historical construct, which in turn raises the question of the origins of gerundive morphology, and ultimately the constructions themselves. The morphology of Truku gerunds suggests that indicative and subjunctive sets, although alike in behavior, originate in two different historical sources. What is striking about these two lines of development is that they share (i) a sort of functional shift with semantic widening, whereby the morphemes ceased to foreground particular thematic roles (resulting forms referring to events), and (ii) the preservation of the original functions of the morphemes, creating a plethora of polysemous forms. These forms would certainly place a heavy load on cognitive processing, with multiple syntactic representations associated with phonetically identical strings of words. Case studies such as this stimulate the need for further empirical investigation into the cognitive foundation for language variation and change.

Finally, the topic relates to the debate over the nature of predication in Philippine-type voice systems. As discussed in Chapter 4, the "voice-marked forms" in Philippine-type languages have been often analyzed as essentially nominal, with the consequence that their sentences are described as copular or equational. By virtue of being between-category entities, the existence of gerunds that fit the definition that I adopted in this dissertation clearly presupposes the existence of the two categories, nouns and verbs, at the syntactic level. The categoriality of predicates in Formosan languages has not been openly questioned in the literature. This study confirms that, as far as Truku is concerned, the conventional assumption of the two syntactic categories is justified. Note, however, that this analysis does not provide any presupposition for precategoriality of lexical roots in such languages, or the lack thereof.

6.3. Limitations and future research

There are a number of limitations that need to be addressed before concluding this dissertation. First, any study of an endangered language is faced with a rapidly decreasing number of proficient native speakers. This study focused on data from a handful of speakers both due to their availability and for the sake of consistency. Ideally, a wider range of speakers should be included in future studies to assess the productivity of the constructions, as well as their

possible loss among younger speakers. Additionally, the nature of the data used in this study was largely limited to elicited sentences. While gerunds are observed fairly frequently in written texts, data from a more naturalistic context would provide additional information as to their actual usage. For instance, the presence/absence of pragmatic cues may well shed light on the level of cognitive load that the polysemous morphology presents.

As for the historical analyses of gerundive morphology, there are prospects for further assessment of their plausibility. The best odds of finding a missing link to strengthen the hypotheses presented in this dissertation certainly lie in Atayal varieties, which are the closest relatives to Seediq. Atayal and Seediq are the only constituents of Atayalic, a primary branch of Austronesian. In this dissertation, most of the extra-Seediq material for the origin of Ce-/Cu- reduplication was drawn from a single Atayal variety: Mayrinax. This was partially motivated by the more comprehensive morphosyntactic description available for Mayrinax compared to other varieties. Moreover, the phonological conservatism of this variety made it the most ideal candidate for discerning historical connections with reduplicative patterns in Seediq varieties. It was found that both Seediq Ce-/Cu- reduplication and Mayrinax Ca- reduplication are used for future tense marking. Meanwhile, some constructions that can possibly be analyzed with manner or event reading in Mayrinax have been reported.

On the other hand, event-denoting functions of affixes like <*n*>, *-an*, and *-un* are thus far unattested in Atayal varieties. I do not know whether such constructions are actually non-existent or simply have been overlooked. A more rigorous search for reconstructed morphological functions at the Proto-Atayal level may result in establishing a more continuous relationship with their proposed original sources in PAn. I hasten to add that the level of diversity among Atayalic varieties poses a challenge to this endeavor.

Lastly, describing the construction at the heart of this dissertation underscored a fundamental challenge in linguistic description and typology. Namely, the existence of a syntactic category cannot be assumed for two languages, let alone comparable types of *between-category* entities. To cite William Croft: "[the] fundamental prerequisite for cross-linguistic comparison is cross-linguistic comparability, that is the ability to identify the same grammatical phenomenon across languages…This is in fact a fundamental issue in all linguistic theory" (2003:13). In Austronesian linguistics, the issue of comparability is perhaps most notably exemplified in the

long-standing controversy over what constitutes the "subject" in Philippine-type languages (see Guilfoyle, Hung, and Travis 1992; Kroeger 1993; Schachter 1976, 1996, among others).

Cross-linguistic search for "gerunds" made it apparent that, not only are they insufficiently described, but there is also an overall lack of agreement in the definition of such a construction (or any other type of nominalization) in the literature. I believe that the lack of attention to gerunds in the literature can be mainly attributed to two reasons: (i) the elusive nature of the construction and (ii) its rarity in natural speech.

As we saw in Chapter 3, nominalizations in general are notoriously difficult to define due to their trans-categorial nature. Often, there are multiple nominalizations within a language that lie on a cline between sentential and nominal. Gerunds, in particular, tend to lie somewhere in the middle of the nominalization continuum, making their cross-linguistic identification even harder. As a result, some authors use the process of elimination whereby they analyze the most nominal of the nominalizations in a language as derived nominals, the most sentential ones as nominalized clauses, and finally apply the label "gerund" to constructions that do not fit either of the two categories. Thus, the constructions that are called "gerunds" do not necessarily share many characteristics across languages except that they are some sort of nominalization.

Furthermore, there is a possibility that gerunds are generally disfavored in natural speech if an alternative strategy is available. Intuitively, native speakers of English seem to prefer sentences like (6.3), with a finite clause as a complement, over ones like (6.4), with a gerund, for encoding the same event.

(6.3) I was disappointed that John refused the offer.
(6.4) I was disappointed by John's refusing the offer.

Although I do not have access to frequency counts for gerunds in Truku, it may well be that their occurrences are relatively rare. If this turns out to be a cross-linguistic tendency, it would surely explain the shortage of description of the construction. In turn, avoidance of gerunds may be a direct result of the in-between category nature of the construction, which would impose a higher cognitive load in processing.

Despite the elusiveness of the construction, I sought, to the best of my capabilities, to adopt an approach that equally weighs form and function for the sake of cross-linguistic comparisons. It is my hope that all of us working to describe the world's languages continue to strive for the best practice in this regard.

REFERENCES

Abney, Steven P. 1987. *The English noun phrase in its sentential aspect*. PhD dissertation, MIT.

Aldridge, Edith. 2002. Nominalization and WH-movement in Seediq and Tagalog. *Language and Linguistics*, 3(2): 393–426.

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. PhD dissertation, Cornell University.

Alexiadou, Artemis. 2001. *Functional structure in nominals: Nominalization and ergativity*. Amsterdam: John Benjamins.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics.* Oxford: Oxford University Press.

Aronoff, Mark. 1976. *Word formation in generative grammar.* Cambridge, MA: The MIT Press.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark. 1999. External possession in Mohawk: Body parts, incorporation, and argument structure. In Payne and Barshi, 293–323.

Baker, Mark, Roberto Aranovich, and Lucia Golluscio. 2005. Two types of syntactic noun incorporation: Noun incorporation in Mapudungun and its typological implications. *Language*, 81(1): 138–176.

Bell, Sarah J. 1976. Cebuano subjects in two frameworks. PhD dissertation, MIT.

Bell, Sarah J. 1983. Advancements and ascensions in Cebuano. In *Studies in Relational Grammar 1*, ed. by David M. Perlmutter, 143–218. Chicago: University of Chicago Press.

Benton, Richard A. 1971. *Pangasinan dictionary*. Honolulu: University of Hawawi‘i Press.

Berman, Ruth. 1982. Dative marking of the affectee role: Data from Modem Hebrew. *Hebrew Annual Review*, 6:35–59.

Blake, Frank R. 1925. *A grammar of the Tagalog language*. American Oriental Series 1. New Haven: American Oriental Society.

Blank, Andreas, and Peter Koch (eds.). 1999. *Historical semantics and cognition.* Cognitive Research 13. Berlin: Mouton de Gruyter.

Bloomfield, Leonard. 1917. Tagalog texts with grammatical analysis. *University of Illinois Studies in Language and Literature*, 3:2–4.

Blust, Robert. 1996. Notes on the semantics of Proto-Austronesian *-an ‘locative’. *NUSA vol.39*: *Voice in Austronesian*, ed. by Marian Klamer, 1–11.

Blust, Robert. 1998. Ca- reduplication and Proto-Austronesian grammar. *Oceanic Linguistics,* 37 (1): 29–64.

Blust, Robert. 1999a. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul J-K Li, 31–94. Taipei: Academia Sinica.

Blust, Robert. 1999b. A note on covert structure: Ca-reduplication in Amis. *Oceanic Linguistics*, 38(1): 168–174.

Blust, Robert. 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics*, 43(1): 73–148.

Blust, Robert. 2013. *The Austronesian languages, revised edition*. Canberra: Pacific Linguistics.

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In *The nature of explanation in linguistic theory,* ed. by John Moore and Maria Polinsky, 31–67. Stanford, CA: CSLI Publicaions.

Borer, Hagit. 2005. *Structuring sense, Volume II: The normal course of events*. Oxford: Oxford University Press.

Broschart, Jurgen. 1997. Why Tongan does it differently: Categorial distinctions in a language without nouns and verbs. *Linguistic Typology*, 1(2): 123–165.

Carrier-Duncan, Jill. 1985. Linking of thematic roles in derivational word formation. *Linguistic Inquiry*, 16(1): 1–34.

Chang, Anna Hsiou-chuan. 2006. A reference grammar of Paiwan. PhD dissertation, Australian National University.

Chang, Henry Yung-li. 1997. Voice, case, and agreement in Seediq and Kavalan. PhD dissertation, National Tsing Hua University.

Chang, Henry Yung-li. 2000a. 噶瑪蘭語參考語法 [Kavalan reference grammar]. Taipei: Yuan-liu Publishing.

Chang, Henry Yung-li. 2000b. 賽德克語參考語法 [Seediq reference grammar]. Taipei: Yuan-liu Publishing.

Chang, Henry Yung-li, and Amy Pei-jung Lee. 2002. Nominalization in Kavalan. *Language and Linguistics*, 3(2): 349–368.

Chen, Cheng-fu. 2008. Aspect and tense in Rukai: Interpretation and interaction. PhD dissertation, The University of Texas Austin.

Chiang, Wen-yu, and Fang-mei Chiang. 2005. Vowel dispersion in Truku. *Concentric: Studies in Linguistics,* 31(1): 21–45.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. by Roderick A. Jacobs and Peter S. Rosenbaum, 184–221. Waltham, MA: Ginn and Company.

Chung, Sandra. 1976. On the Subject of two passives in Indonesian. In *Subject and Topic*, ed. by Charles. N. Li, 57–98. New York: Academic Press.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Comrie, Bernard, and Sandra A. Thompson. 1985. Lexical nominalization. In *Language typology and syntactic description, volume III: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 349–398. Cambridge: Cambridge University Press.

Comrie, Bernard, and Sandra A. Thompson. 2007. Lexical nominalization. In *Language typology and syntactic description, volume III: Grammatical categories and the lexicon*, *2nd edition*, ed. by Timothy Shopen, 334–381. Cambridge: Cambridge University Press.

Cooreman, Ann, Barbara A. Fox, and Talmy Givón. 1984. The discourse definition of ergativity. *Studies in Language*, 8(1): 1–34.

Council of Indigenous Peoples. n.d. http://www.apc.gov.tw/.

Council of Indigenous Peoples. n.d. Yuanzhuminzu yuyan xianshang cidian (原住民族語言線上詞典) [Online dictionaries of aboriginal languages]. http://e-dictionary.apc.gov.tw/.

Croft, William. 1985. Indirect object "lowering." *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, ed. by Mary Niepokuj, Mary VanClay, Vassiliki Nikiforidou, and Deborah Feder, 39–51. Berkeley, CA: Berkeley Linguistics Society.

Croft, William. 2003. *Typology and universals, 2nd edition.* Cambridge: Cambridge University Press.

Dahl, Otto Christian. 1973. *Proto-Austronesian.* Scandinavian Institute of Asian Studies Monograph Series no. 15. Lund: Studentlitteratur.

De Wolf, Charles. 1988. Voice in Austronesian languages of Philippine type: Passive, ergative, or neither? In *Passive and voice*, ed. by Masayoshi Shibatani, 143–193. Amsterdam: John Benjamins.

Deal, Amy Rose. 2013. External possession and possessor raising. To appear in *The Blackwell companion to syntax, 2nd edition*, ed. by Martin Everaert, and Henk C. van Riemsdijk. Hoboken, NY: Wiley-Blackwell.

Dixon, R. M. W. 1979. Ergativity. *Language*, 55(1): 59–138.

Donohue, Mark. 1999. Syntactic roles vs. semantic roles: External possession in Tukang Besi. In Payne and Barshi, 373–401.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language*, 67(3): 547–619.

Egerod, Søren. 1966. A statement on Atayal phonology. *Artibus Asiae Supplementum Vol.23, Essays offered to G. H. Luce by his colleagues and friends in honour of his seventy-fifth birthday. Vol. 1: Papers on Asian history, religion, languages, literature, music folklore, and anthropology*, ed. by Jean Boisselier, Alexander B. Griswold, Ba Shin, and Gordon H. Luce, 120–130. Zurich: Artibus Asiae Publishers.

Ferrell, Raleigh. 1969. *Taiwan aboriginal groups: Problems in cultural and linguistic classification*. Institute of Ethnology, Academia Sinica, Monograph no. 17. Taipei: Academia Sinica.

Foley, William, and Robert D. Van Valin 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Foley, William. 2007. The place of Philippine languages in a typology of voice systems. In *Voice and grammatical relations in Austronesian languages*, ed. by Simon Musgrave and Peter Austin, 22–44. Stanford, CA: CSLI Publications.

Fox, Barbara A. 1981. Body part syntax: Towards a universal characterization. *Studies in Language*, 5(3): 323–342.

Gerdts, Donna B. 1988. Antipassives and causatives in Ilokano. In *Studies in Austronesian Linguistics*, ed. by Richard McGinn, 295–231. Athens, OH: Ohio University Center for International Studies.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Grimshaw, Jane. 1990. *Argument structure.* Cambridge, MA: The MIT Press.

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. SPEC of IP and SPEC of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory,* 10(3): 375–414.

Harley, Heidi. 2009. The morphology of nominalizations and the syntax of *v*P. In *Quantification, definiteness and nominalization*, ed. by Monika Rathert and Anastasia Giannadikou, 320–342. Oxford: Oxford University Press.

Harley, Heidi, and Rolf Noyer. 1998. Mixed nominalizations, short verb movement and object shift in English. *Proceedings of North East Linguistic Society 28*, ed. by Pius N. Tamanji and Kiyomi Kusumoto, 143–157. Amherst, MA: GLSA.

Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language*, 22(2): 315–351.

Haspelmath, Martin. 1999. External possession in a European areal perspective. In Payne and Barshi, 109–135.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Himmelmann, Nikolaus P. 1991. The Philippine challenge to universal grammar. Arbeitspapier nr. 15 (Neue Folge). Köln: Institut für Sprachwissenschaft der Universität zu Köln.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. *The Austronesian languages of Asia and Madagascar*, ed. by Alexander Adelaar and Nikolaus P. Himmelmann, 110–181. New York: Routledge.

Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In *Voice and grammatical relations in Austronesian languages*, ed. by Simon Musgrave and Peter Austin, 247–293. Stanford, CA: CSLI Publications.

Holmer, Arthur. 1996. *A parametric grammar of Seediq*. Lund: Lund University Press.

Holmer, Arthur. 2002. The morphology and syntax of Seediq focus. In *The history and typology of Western Austronesian voice systems*, ed. by Fay Wouk and Malcolm Ross, 333–354. Canberra: Pacific Linguistics.

Holmer, Arthur, and Loren Billings. 2014. Clitic pronouns in Seediq. *Argument realisations and related constructions in Austronesian languages: Papers from 12-ICAL, volume 2*, ed. by I Wayan Arka and N.L. L. Mas Indrawati, 111–139. Canberra: Asia-Pacific Linguistics.

Hsieh, Fuhui. 2011. The functions of *-an* and *=ay* in Kavalan. In *Nominalization in Asian languages*, ed. by Yap, Foong Ha, Grunow-Hårsta, Karen, Wrona, Janick, 501–522. Amsterdam: John Benjamins.

Huang, Hui-chuan J. 2006. Squliq Atayal syllable onset: Simple or complex? *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, ed. by Da-an Ho, Henry Y. Chang, and Lillian M. Huang, 489–505. Language and Linguistics Monograph Series W-5. Taipei: Academia Sinica.

Huang, Lillian M. 2000. 泰雅語參考語法 [Atayal reference grammar]. Taipei: Yuan-liu Publishing.

Huang, Lillian M. 2002. Nominalization in Mayrinax Atayal. *Language and Linguistics*, 3(2): 197–225.

Imanishi, Kazuhiro. 2009. A basic description of the Amis language (Taiwan). MA thesis, University of Tokyo.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language*, 51(3): 639–671.

Jespersen, Otto. 1940. *A modern English grammar on historical principles, part IV: Syntax.* London: G. Allen and Unwin.

Kaufman, Daniel. 2009. Austronesian Nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics*, 35(1): 1–49.

Kiyomi, Setsuko. 1993. A typological study of reduplication as a morpho-semantic process: Evidence from five language families (Bantu, Australian, Papuan, Austroasiatic and Malayo-Polynesian). PhD dissertation, Indiana University.

Kratzer, Angelika. 1996. Severing the external argument from its verb. *Phrase structure and the lexicon*, ed. by Johan Rooryck and Laurie Zaring, 109–137. Dordrecht: Springer.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.

Kuo, Jonathan Cheng-chuen. 2014. The origin of marked order in Seediq clitic pronoun clusters: A historical/typological perspective. Unpublished ms.

Langacker, Ronald W. 1993. Reference-point constructions. *Cognitive Linguistics*, 4(1): 1–38.

Lee, Amy Pei-jung. 2012. The Segment /w/ and contrastive hierarchy in Paiwan and Seediq. *Concentric: Studies in Linguistics*, 38(1): 1–37.

Li, Paul Jen-kuei. 1980. The phonological rules of Atayal dialects. *Bulletin of the Institute of History and Philology, Academia Sincia*, 51(2): 349–405.

Li, Paul Jen-kuei. 1981. Reconstruction of Proto-Atayalic phonology. *Bulletin of the Institute of History and Philology, Academia Sincia*, 52(2): 235–301.

Li, Paul Jen-kuei. 2002. Nominalization in Pazih. *Language and Linguistics*, 3(2): 227–239.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2006. *Kavalan dictionary*. Language and Linguistics Monograph Series A-19. Taipei: Institute of Linguistics, Academia Sincia.

Lieber, Rochelle. 1980. On the organization of the lexicon. PhD dissertation, MIT.

Lin, Hui-shan. 2015. Squliq Atayal reduplication: Bare consonant or full syllable copying? In *New advances in Formosan linguistics*, ed. by Elizabeth Zeitoun, Stacy F. Teng, and Joy J. Wu, 75–100. Canberra: Pacific Linguistics.

Liu, Tsai-hsiu. 2011. Complementation in three Formosan languages —Amis, Mayrinax Atayal and Tsou. PhD dissertation, University of Hawai‘i at Mānoa.

Malouf, Robert P. 1998. Mixed categories in the hierarchical lexicon. PhD dissertation, Stanford University.

McGregor, William. 1999. External possession in Nyulnyulan languages. In Payne and Barshi, 429–448.

Moravcsik, Edith A. 1978. Reduplicative constructions. In *Universals of human language vol. 3: Word structure*, ed. by Joseph H. Greenburg, Charles A Ferguson, and Edith A. Moravcsik, 297–334. Stanford, CA: Stanford University Press.

Mosel, Ulrike, and Even Hovdhaugen. 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.

Naylor, Paz B. 1980. Linking, relation-marking, and Tagalog syntax. In *Austronesian studies: Papers from the Second Eastern Conference on Austronesian Languages*, ed. by Paz B. Naylor, 33–49. Ann Arbor: The University of Michigan.

Niepokuj, Mary. 1997. *The development of verbal reduplication in Indo-European*. Journal of Indo-European studies monograph no. 24. Washington DC: Institute for the Study of Man.

Ochiai, Izumi. 2009. Tgadaya Seediq pronoun and argument realizations. MA thesis, National Taiwan University.

Oiwa, Mayumi. 2016. Possessor raising in Truku Seediq. In *AFLA 23: The Proceedings of the 23rd Meeting of the Austronesian Formal Linguistics Association*, ed. by Hiroki Nomoto, Takuya Miyauchi, and Asako Shiohara, 194–209. Canberra: Asia-Pacific Linguistics.

Payne, Doris L., and Immanuel Barshi (eds). 1999. *External possession*. Amsterdam: John Benjamins.

Payne, Thomas E. 1982. Role and Reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language*, 6(1): 75–106.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: The MIT Press.

Rau, Der-hwa Victoria. 1992. A grammar of Atayal. PhD dissertation, Cornell University.

Rau, Der-hwa Victoria. 2002. Nominalization in Yami. *Language and Linguistics*, 3(2): 165–195.

Reid, Lawrence. 2009. On the diachronic development of C_1V_1- reduplication in some Austronesian languages. *Morphology*, 19(2): 239–261.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In *The history and typology of Western Austronesian voice systems*, ed. by Fay Wouk and Malcolm Ross, 17–62. Canberra: Pacific Linguistics.

Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 295–326. Canberra: Pacific Linguistics.

Schachter, Paul. 1976. The subject in Philippine languages: Topic, Actor, Actor-Topic, or none of the above? In *Subject and topic*, ed. by Charles N. Li, 491–518. New York: Academic Press.

Schachter, Paul. 1996. The subject in Tagalog: Still none of the above. UCLA Occasional Papers in Linguistics no. 15.

Shibatani, Masayoshi. 1988. Voice in Philippine languages. *Passive and voice*, ed. by Masayoshi Shibatani, 85–142. Amsterdam: John Benjamins.

Shibatani, Masayoshi. 1994. An integrational approach to possessor raising, ethical datives, and adversative passives. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session Dedicated to the Contributions of Charles J. Fillmore*, ed. by Susanne Gahl, Andy Dolbey, and Christopher Johnson, 461–486. Berkeley, CA: Berkeley Linguistics Society.

Soyang Patas. [Holy Bible]. 1988. The Bible Society in the Republic of China.

Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid. 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics, vol. 2, Tracking the travellers*, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 145–170. Canberra: Pacific Linguistics.

Štekauer, Pavol, Salvador Valera, and Lívia Körtvélyessy. 2012. *Word-formation in the world's languages: A typological survey*. Cambridge: Cambridge University Press.

Sung, Li-May. 2011. Clausal nominalization in Budai Rukai. In *Nominalization in Asian languages*, ed. by Foong Ha Yap, Karen Grunow-Hårsta, and Janick Wrona, 523–559. Amsterdam: John Benjamins.

Taiwan Indigenous People's Knowledge Economic Development Association. n.d. http://www.twedance.org/.

Tang, Apay Ai-yu. 2011. From diagnosis to remedial plan: A psycholinguistic assessment of language shift, L1 proficiency, and language planning in Truku Seediq. PhD dissertation, University of Hawai‘i at Mānoa.

Tang, Chih-chen Jane. 2002. On nominalizations in Paiwan. *Language and Linguistics,* 3(2): 283–333.

Teng, Stacy Fang-ching. 2007. A reference grammar of Puyuma, an Austronesian language of Taiwan. PhD dissertation, Australian National University.

Tsukida, Naomi. 2009. Sedekku-go no bumpou [A grammar of Seediq]. PhD dissertation, Tokyo University.

Tuggy, David. 1980. Ethical dative and possessor omission Si, possessor ascension No!. In *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session,* ed. by John Daly P. and Margaret H. Daly, 24:97–141.

Ura, Hiroyuki. 1996. Multiple feature-checking: A theory of grammatical function splitting. PhD dissertation, MIT.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review*, 66(2): 143–160.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Wasow, Thomas, and Thomas Roeper. 1972. On the subject of gerunds. *Foundations of Language*, 8(1): 44–61.

Wechsler, Stephen, and I Wayan Arka. 1998. Syntactic ergativity in Balinese: An argument structure based theory. *Natural Language and Linguistic Theory*, 16(2): 387–441.

Wolff, John U. 1973. Verbal inflection in Proto-Austronesian. *Parangal Kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday*, ed. by Andrew Gonzalez, 71–94. Quezon City: Linguistic Society of the Philippines.

Wu, Jing-lan Joy. 2006. Verb classification, case marking, and grammatical relations in Amis. PhD dissertation, SUNY Buffalo.

Yeh, Marie Mei-li. 2006. The grammaticalization of *’am* in Saisiyat’. *Streams converging into an ocean, Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, ed. by Da-an Ho, Henry Y. Chang, and Lillian M. Huang, 393–413. Language and Linguistics Monograph Series W-5. Taipei: Institute of Linguistics, Academia Sinica.

Zeitoun, Elizabeth. 2002. Nominalization in Mantauran (Rukai). *Language and Linguistics*, 3(2): 241–282.

Zeitoun, Elizabeth. 2007. *A grammar of Mantauran (Rukai).* Taipei: Institute of Linguistics, Academia Sinica.

Zeitoun, Elizabeth, Tai-hwa Chu, and Lalo a tahesh kaybaybaw. 2015. *A study of Saisiyat morphology.* Oceanic Linguistics Special Publication no. 40. Honolulu: University of Hawai‘i Press.

Zeitoun, Elizabeth, and Lillian M. Huang. 2000. Concerning *ka-*, an overlooked marker of verbal derivation Formosan languages. *Oceanic Linguistics,* 39(2): 391–414.

Zeitoun, Elizabeth, Lillian M. Huang, Marie M. Yeh, Anna H. Chang, and Joy Y. Wu. 1996. The temporal, aspectual, and modal systems of some Formosan languages: A typological perspective. *Oceanic Linguistics,* 35(1): 21–56.

Zeitoun, Elizabeth, and Chen-huei Wu. 2006. An overview of reduplication in Formosan languages. *Streams converging into an ocean, Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, ed. by Da-an Ho, Henry Y. Chang, and Lillian M. Huang, 97–

142. Language and Linguistics Monograph Series W-5. Taipei: Institute of Linguistics, Academia Sinica.